Phenotypic and genetic characterisation of an extended pustular psoriasis dataset

Twelves, Sophie Maria

Awarding institution:
King's College London

PHENOTYPIC AND GENETIC CHARACTERISATION OF AN EXTENDED PUSTULAR PSORIASIS DATASET

By

Sophie Maria Twelves

A thesis submitted in fulfilment of the requirements for the degree of

Doctor of Philosophy at King's College London

Department of Medical and Molecular Genetics

King's College London

9th floor, Tower Wing, Guy's Hospital

Great Maze Pond, London SE1 9RT

July 2019

ACKNOWLEDGEMENTS

First, I would like to express my deepest gratitude for the support, guidance and knowledge provided by my primary supervisor, Dr Francesca Capon. She has been fundamental to the progress I have made during the past three years, both inside and outside of science. Her commitment to reviewing drafts of my thesis and helping me prepare for my viva was invaluable and I am extremely appreciative of the time she has spent on this.

I also want to thank my second supervisor, Professor Jonathan Barker. He has been instrumental in facilitating many aspects of my work and also supported my attendance of a number of excellent conferences. Both he and Professor Catherine Smith have developed a positive and exciting environment surrounding psoriasis research within St John's Institute of Dermatology, which I feel lucky to have been a part of.

Professor Rebecca Oakey, Professor David Kelsell and Dr Myles Lewis have been crucial to the development of my PhD work through their roles on my PhD progress committee. I am grateful for their commitment, interest and constructive feedback.

I would also like to thank Professor Oakey and the rest of the Kings Biosciences Institute MRes/PhD programme team for their support over the past four years. In addition, I wish to acknowledge the Guy's and St Thomas' Charity Trust for funding my MRes and PhD.

I am appreciative of the work of Professor Alex Navarini, the ERASPEN consortium, those involved in APRICOT and PLUM, and the numerous other investigators who have recruited the patients analysed in this study. Rosemary Wilson and Dr Alshimaa Mostafa were instrumental in the compilation of the final dataset, and Dr Nick Dand provided extremely helpful statistical advice.

Special thanks must go to the former and current members of the Capon group, particularly to Satveer, Marta and Marika for being such kind, funny and knowledgeable colleagues. Without

them my days in the lab would have been much less enjoyable. It has also been a pleasure to get to know Karoliina, Thanos and Natashia over the past 18 months.

I find it difficult to express how lucky I have been to work with such a wonderful group of people on the 9th floor, both past and present. We have shared the highs and lows of both our work and personal lives, and the support and friendship I have experienced will stay with me forever. I will remember our wide-ranging and thought-provoking lunchtime discussions, and our evenings spent in The Miller!

I particularly want to mention Giulia and Heather, who have been the best office bay companions one could wish for: a source of advice, laughter and information about free food. And, most of all, Rosie, who not only was always happy to help or teach me in the lab, but who has also become one of my closest friends.

Outside of KCL, thank you to Emily, Eric, Ollie, Grace, Jin-Xi, Rob, Robyn, Rachel, Sarah, Seb and all my other friends for always lifting my spirits and making me smile. And for being so understanding when I have repeatedly forgotten respond to texts over the past few months!

I am grateful for the kindness of my partner's family. Day trips, dinners and cinema visits with them provided much-needed injections of normality and fun when the experiments weren't behaving!

My extended family have been a constant source of encouragement and affection. Thank you particularly to my brother, Andrew, for never taking me too seriously, to my Dad for the many hours of advice and reassurance, and to my Mum for always being there when I needed her.

Last but certainly not least, I would like to thank my partner, Finn. He has helped me through the difficult periods, celebrated my successes and, perhaps most impressively, put up with my mood swings and absent mindedness while I wrote this thesis! I find it hard to imagine reaching this point without him.

CONTRIBUTIONS

I hereby declare the work presented in this thesis was performed exclusively by me, with the following exceptions. Elias Burri and Dr Alshimaa Mostafa (University Hospital Zurich and Beni Suef University, Egypt) compiled genetic and phenotypic data for patients in the ERASPEN database. Dr Asma Smahi (Necker-Enfants Malades Hospital, France) provided three *IL36RN* constructs and modified HeLa cells were provided by Professor Seamus Martin (The Smurfit Institute, Trinity College, Dublin). Within the Capon group, Karoliina Hassi produced four of the *IL36RN* constructs and carried out the Western blot assay in HeLa cells. Athanasios Niaouris contributed to the screening of *ZNF33A* in patients. The whole-exome sequencing experiment was undertaken by technical staff from the Genomics Core facility of Guy's and St Thomas' Hospital Biomedical Research Centre (BRC). The raw sequence data was processed by the BRC Bioinformatics Core.

ABSTRACT

Autoinflammatory diseases are a group of conditions (often monogenic) that are characterised by episodes of inflammation driven by the innate immune system. One example is pustular psoriasis, a rare and severe disorder that manifests with the eruption of sterile pustules on erythematous skin. The condition has been classified into three subtypes, (generalised pustular psoriasis, acrodermatitis continua of Hallopeau and palmoplantar pustulosis), which can be distinguished based on the localisation and extent of skin lesions, and on the presence or absence of general symptoms.

Previous genetic studies have identified variants in *IL36RN*, *AP1S3* and *CARD14* in a proportion of affected individuals. Such patients, however, account for fewer than 30% of disease cases. Moreover, the growing use of *IL36RN* sequencing as a diagnostic tool has raised issues relating to the interpretation of discovered missense variants.

In this context, the aim of the project was to improve our understanding of pustular psoriasis genetics and phenotypic variability.

In the first stage of the study, the analysis of an extended patient cohort (n=863) revealed clinical, demographic and genetic factors that differentiated palmoplantar pustulosis from the other forms of the disease. A significant (P = 0.003), dose-dependent association between *IL36RN* status and age of onset was also observed.

In the second part of the research, a set of *in-vitro* assays were designed to measure the impact of *IL36RN* missense changes on protein function. Comparison with predictions generated by online tools demonstrated the classification errors which can arise when *in-silico* methods alone are utilised to assess missense changes.

Finally, the study sought to identify new genetic determinants for pustular psoriasis, by the means of whole-exome sequencing. Two candidate genes were uncovered and followed-up in a larger dataset. This highlighted the genetic heterogeneity of the condition as well as the importance of analysing representative control populations.

Taken together, these findings have improved our understanding of known pustular psoriasis genes and paved the way for the identification of novel genetic determinants for the disease.

Table of Contents

List of Tables

List of figures

Abbreviations

ACH	Acrodermatitis continua of Hallopeau
ACMG	American College of Medical Genetics
AGEP	Acute generalized exanthematous pustulosis
AhR	Aryl hydrocarbon receptor
AHRR	Aryl hydrocarbon receptor repressor
ALS	Amyotrophic lateral sclerosis
AMP	Association for Molecular Pathology
AP-1	Adapter related protein complex 1
AP1S3	Adaptor related protein complex 1 subunit sigma 3 gene
APRICOT	Anakinra for Pustular psoriasis: Response in a Controlled Trial
APσ1C	Adaptor related protein complex 1 subunit sigma 3
ARFGAP2	ADP ribosylation factor GTPase activating protein 2
ATAC-seq	Assay for transposase-accessible chromatin using sequencing
B2M	Beta-2 microglobulin
BLAST	Basic Local Alignment Search Tool
BRC	Biomedical Research Council
BRCA2	BRCA2, DNA repair associated
C21orf2	Chromosome 21 open reading frame 2
CADD	Combined Annotation Dependent Depletion
CARD	Caspase recruitment domain
CCL20	C-C motif chemokine ligand 20
cDNA	Complementary DNA
ChIP-seq	Chromatin immunoprecipitation sequencing
CMV	Cytomegalovirus
CNV(s)	Copy number variant(s)
COL6A1	Collagen type VI alpha 1 chain
CRF	Case report form
Ct	Threshold cycle

DBD	DNA-binding domain
dbNSFP	Database for non-synonymous functional predictions
DC	Dendritic cell
ddH_2O	Double distilled water
DEG	Differentially expressed genes
DMEM	Dulbecco's Modified Eagle's Medium
DNase	Deoxyribonuclease
dNTPs	Deoxyribonucleotide triphosphate
ELISA	Enzyme-linked immunosorbent assay
ERASPEN	European Rare and Severe Psoriasis Expert Network
ESP	Exome Sequencing Project
ESR	Erythrocyte sedimentation rate
EVI5	Ecotropic viral integration site 5
EWAS	Exome-wide association study
ExAC	Exome Aggregation Consortium
FAP	Functional annotation of psoriasis susceptibility alleles
FBS	Foetal bovine serum
GERP	Genomic Evolutionary Rate Profiling
gnomAD	Genome Aggregation Database
GPP	Generalised pustular psoriasis
GRR	Genetic risk ratio
GWAS	Genome-wide association studies
HDR	Homology-directed repair
HEK293	Human embryonic kidney 293 (cell line)
HeLa	Henrietta Lacks (cell line)
HGMD	Human gene mutation database
HLA	Human leukocyte antigen
IGV	Interactive genome viewer
IL	Interleukin
IL-1RAcP	Interleukin 1 receptor accessory protein

IL1RL2	Interleukin-1 receptor-like 2
IL36A	Interleukin 36 alpha gene
IL-36R	Interleukin 36 receptor
IL-36Ra	Interleukin 36 receptor antagonist
IL36RN	Interleukin 36 receptor antagonist gene
INDEL(s)	Insertion(s) or deletion(s)
IRAK	Interleukin 1 receptor associated kinase
IκB	Inhibitor of nuclear factor kappa B kinase
JADE1	Jade family PHD finger 1
KAP1	KRAB-associated protein 1
KC	Keratinocytes
KRAB	Krüppel associated box
LB	Lysogeny broth
LOF	Loss of function
MAF	Minor allele frequency
MALP-2	Macrophage-activating lipopeptide 2
MAP	Mitogen activated protein
MAPK	Mitogen activated protein kinase
MH	Mantel Haenszel
MyD88	Myeloid differentiation primary response 88
nAChR(s)	Nicotinic acetylcholine receptor(s)
NEB	Nebulin
NF-κB	Nuclear factor kappa B
NGS	Next generation sequencing
NPHP3	Nephrocystin 3
p62	Nucleoporin 62
PAH	Pulmonary arterial hypertension
PBMCs	Peripheral blood mononuclear cells
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction

PKC	Primary keratinocytes
PLUM	Pustular psoriasis, eLucidating Underlying Mechanisms
PPARG	Peroxisome proliferator activated receptor gamma
PPIA	Peptidylprolyl isomerase A
PPP	Palmoplantar pustulosis
PROVEAN	Protein Variation Effect Analyzer
PRP	Pityriasis rubra pilaris
PRTFDC1	Phosphoribosyl transferase domain containing 1
PSI-BLAST	Position-specific iterated Basic Local Alignment Search Tool
PUVA	Psoralen + ultraviolet A
PVDF	Polyvinylidene fluoride
qPCR	Quantitative polymerase chain reaction
RNA-seq	RNA sequencing
RPKM	Reads per kilobase per million
RPMI medium	Roswell Park Memorial Institute medium
RQ	Relative quantification
S100A7	S100 calcium binding protein A7
SD	Standard deviation
SDS	Sodium dodecyl sulphate
SIFT	Sorting Intolerant From Tolerant
SLC45A4	Solute carrier family 45 member 4
SLE	Systemic lupus erythematosus
SLR	Synthetic long read
SMRT	Single molecule real time
SNV(s)	Single nucleotide variant(s)
SPAG5	Sperm associated antigen 5
SSMP	Singapore sequencing Malay project
TAB1	TGF-Beta activated kinase 1 (MAP3K7) binding protein 1
TAK1	Mitogen-activated protein kinase 7
TBE	Tris-borate-EDTA

TBS	Tris-buffered saline
TE	Tris-EDTA
Th17	T-helper cell type 17
TIRF	Total internal reflection fluorescence
TLN2	Talin 2
TLR-2/6	Toll-like receptor 2/6
Tm	Melting temperature
TNFA	Tumour necrosis factor alpha gene
TNFα	Tumour necrosis factor alpha
Tollip	Toll interacting protein
TRAF6	TNF receptor associated factor 6
Tris-HCL	Tris-hydrochloric acid
VGLL3	Vestigial like family member 3
VUS	Variant of unknown significance
WES	Whole exome sequencing
WGS	Whole genome sequencing
WWOX	WW domain containing oxidoreductase
ZNF	Zinc finger

1 Introduction

1.1 Pustular psoriasis

The term pustular psoriasis encompasses a group of rare and severe neutrophilic skin disorders with overlapping clinical and genetic features. Consensus diagnostic criteria were lacking until recently, but the European Rare and Severe Psoriasis Expert Network (ERASPEN) has sought to address this issue (1).

1.1.1 Clinical presentation

All forms of pustular psoriasis are characterised by the appearance of macroscopic neutrophil-filled pustules on the skin. These are non-infectious and form as a primary lesion, distinguishing the condition from those where pustules are infected (e.g. bullous impetigo) or form as a secondary feature (e.g. pustules within pre-existing psoriasis plaques) (1).

Three subtypes of pustular psoriasis can be defined based on the location of pustules and the disease course (Figure 1.1, panels a-c). These are generalised pustular psoriasis (GPP), acrodermatitis continua of Hallopeau (ACH) and palmoplantar pustulosis (PPP, also known as palmoplantar pustular psoriasis).

In all three variants, histological staining of affected skin shows neutrophil infiltration into necrotic epidermis, forming spongiform pustules of Kogoj (2) (Figure 1.1, panel d). These are accompanied by intraepidermal micropustules, parakeratosis and elongation of rete ridges (3,4). As pustules dry they may form brown scabs which gradually slough off, with the potential for fresh pustules to appear on the skin beneath (1,3).

Patients with pustular psoriasis may also suffer from plaque psoriasis. This is a more common and less severe skin disease that presents with well demarcated scaly red plaques. It affects 2-5% of the world population, although the prevalence identified in national studies varies widely (3,5).

1.1.1.1 Generalised pustular psoriasis

According to the ERASPEN criteria, generalised pustular psoriasis can be diagnosed based on the observation of macroscopically visible, sterile pustules on non-acral skin. These lesions, which typically appear on erythematous skin, may then expand and coalesce to form lakes of superficial pus (1).

Affected individuals can experience a relapsing (>1 episode) or persistent (>3 months) course of disease. In the former, flares are often accompanied by systemic inflammation, with patients exhibiting malaise, fever, neutrophilia, elevated levels of C-reactive protein, a high erythrocyte sedimentation rate (ESR) and neutrophilic cholangitis (6–8). This may result in intensive care admission for treatment of dehydration, secondary infections and monitoring of organ function (1,3). In addition, patients may continue to experience symptoms such as chronic skin lesions and fatigue between flares, blurring the line somewhat between 'relapsing' and 'persistent' patients (9). The condition is therefore responsible for significant morbidity and, if systemic inflammation is not managed, it is potentially fatal (10,11).

Data on the prevalence of generalised pustular psoriasis is limited, but reports have ranged from 2 to 7 cases per million (4,12). Incidence is estimated at around 0.6-0.7/million (2,4). While the disease certainly has a genetic component (section 1.1.3), flares can be triggered by a variety of environmental factors. These include infections, pregnancy, drug exposure, stress, hypocalcaemia and withdrawal of corticosteroid treatment. Pustular flares that occur solely during pregnancy and often resolve rapidly after birth were previously known as impetigo

herpetiformis (1,12,13), now simply referred to as 'GPP in pregnancy'. This is a particularly dangerous presentation of GPP, threatening the life of both mother and foetus (8).

It has been reported that the presence or absence of a history of plaque psoriasis can define two subsets of GPP patients with differing genetic backgrounds, mean age of onset and environmental triggers (14,15).

The presentation of GPP is very similar to that of acute generalised exanthematous pustulosis (AGEP). This is an adverse drug reaction presenting with a pustular flare and signs of systemic inflammation after, generally within 48hrs of exposure to certain medications (common triggers include aminopenicillins, sulphonamides, quinolones and pristinamycin). The flare resolves after the culprit drug is withdrawn and, crucially, does not reoccur in the absence of the trigger (16). Thus, the lack of spontaneous relapse in AGEP differentiates it from GPP (1).

1.1.1.2 Localised forms of pustular psoriasis

The rarest form of pustular psoriasis is acrodermatitis continua of Hallopeau, where pustulation is localised to the distal regions of fingers and toes, primarily affecting the nail apparatus (1). The disease may also result in loss of nails, destruction of the nail apparatus and osteolysis of the distal phalanx. It is a chronic (>3 months), slowly developing condition which causes severe pain and morbidity, resulting in functional and occupational disability. Prevalence and incidence data are not available due to the rarity of the disease (17).

Palmoplantar pustulosis is also a chronic and localised disease, but here the pustulation is restricted to the palms and/or soles (1). Affected skin also exhibits erythema, hyperkeratosis and scaling (17). Pustulation may also spread to the digits, but a secondary diagnosis of ACH is only made if the nail apparatus becomes involved (1). As in the case of ACH, the condition has a significant effect on the quality of life of patients but is not life-threatening.

The prevalence of palmoplantar pustulosis is higher than that of generalised pustular psoriasis. The condition is also more common in Asia than Europe, with a reported prevalence of 1200/million in Japan verses estimates of 100-500 cases/million in populations of European descent (18–20). Patients are predominantly female, with onset most common in middle-age (8,17).

Some of the environmental factors contributing to the development of palmoplantar pustulosis are the same as those seen in generalised pustular psoriasis: infections, stress and drug exposure. However, other triggers include trauma, metal sensitivity and thyroid dysfunction. One of the most widely recognised risk factors for PPP is smoking, with many studies finding a high or very high prevalence of the behaviour amongst patients. Symptoms also improve upon cessation of smoking (17,21).

Arthralgia and arthritis are relatively common extracutaneous co-morbidities and in some cases may be present in the context of Sonozaki syndrome or SAPHO (synovitis, acne, palmoplantar pustulosis, hyperostosis, osteitis) syndrome (8,21).

Historically, palmoplantar pustulosis has been viewed as a subtype of plaque psoriasis, as patients often present with both conditions and/or a family history of plaque psoriasis and psoriatic arthritis (21). However, more recent genetic and demographic studies have revealed clear differences distinguishing the two (22).

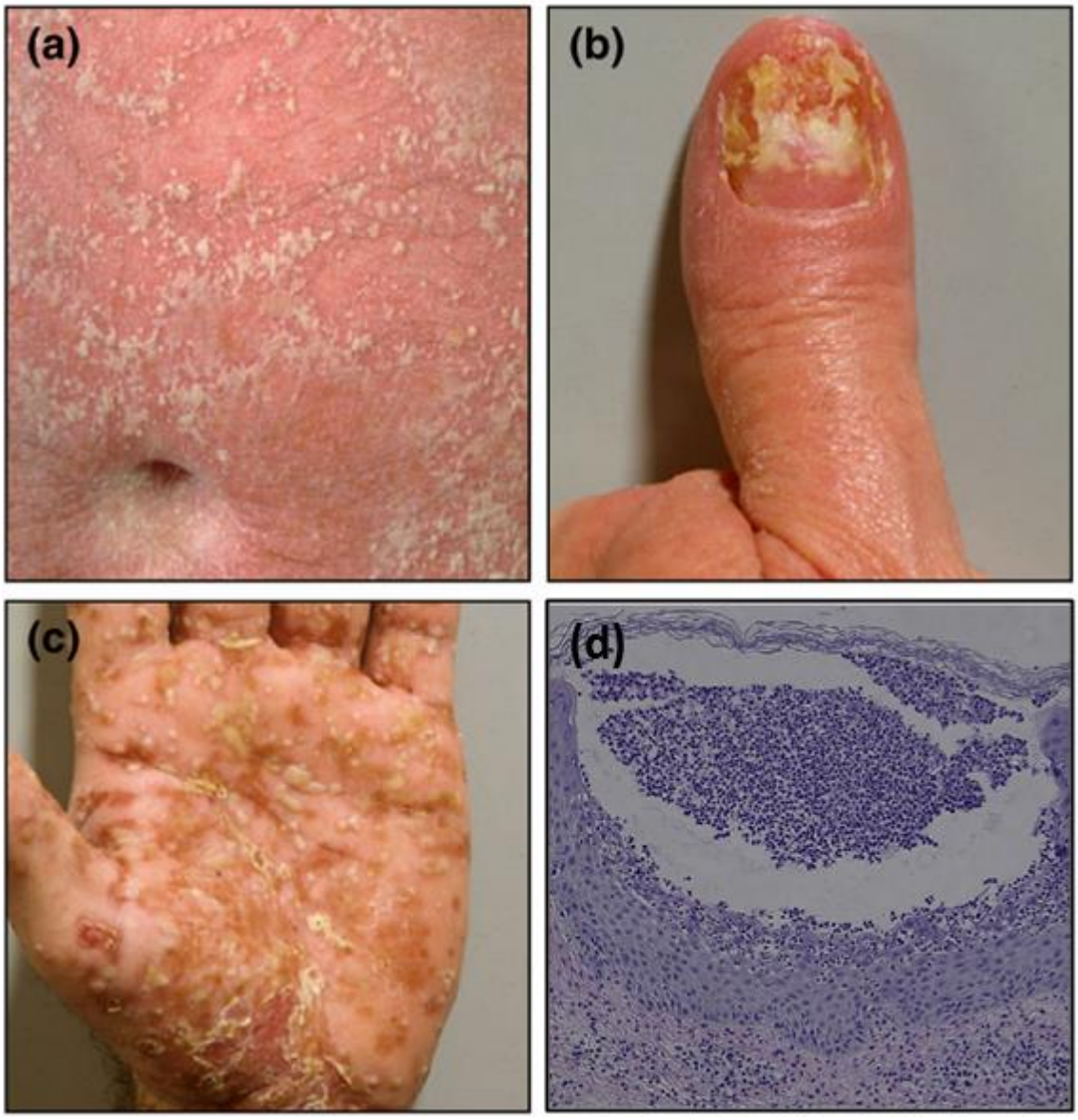


Figure 1.1 Clinical and histological presentation of pustular psoriasis

The figure illustrates the widespread pustulation seen in generalised pustular psoriasis (GPP, panel a), involvement of nail apparatus seen in acrodermatitis continua of Hallopeau (ACH, panel b) and palm lesions seen in palmoplantar pustulosis (PPP, panel c). Panel d shows a histological cross-section of affected epidermis, showing neutrophil infiltration and spongiform pustule of Kogoj. Panels a-c originally published in (1), images obtained with written consent, panel d from (2).

1.1.2 Treatment

Pustular forms of psoriasis are extremely difficult to treat. This is due, at least in part, to the rarity of the disease and the resulting lack of research into potential therapies. While cases are typically managed with medications that are used to good effect in plaque psoriasis, patient response is variable as these agents have a much lower efficacy in the pustular forms of the disease (9,17).

The severe nature of GPP flares necessitates systemic and fast-acting treatment. The common choices for first-line therapy are therefore acitretin, methotrexate, cyclosporin or the TNFα inhibitor infliximab, with the latter two predominating in the most serious cases due to their rapid onset of effect. Second-line treatments include PUVA (psoralen plus ultraviolet-A) phototherapy, topical corticosteroids and TNFα blockers (e.g. etanercept, adalimumab). A combination of first-line oral systemics with second-line biologics may also be used, such as cyclosporin and etanercept or methotrexate and infliximab (17,23).

While systemic steroids are usually avoided in treatment of GPP, due to the risk of relapse upon drug cessation, they can be prescribed when the disease presents during pregnancy. Labour may also be induced, as symptoms often improve rapidly after delivery. Acitretin, in contrast, should not be used in pregnant women as it can disrupt development of the foetus and cause birth abnormalities (17).

More recently developed biological agents targeting IL-1 (anakinra), IL-12/23 (ustekinumab) or IL-17A (secukinumab) have been used to successfully treat some GPP patients. Here the literature is more substantial (6), enabling Boehner et al. to review case series, case reports and clinical trials. Their study found evidence for a therapeutic potential of biologics, with the most robust data obtained for IL-17A blockers (9). These conclusions are however limited by small

patient numbers, the single arm/open label design of trials and the self-remitting course of GPP, and so must be interpreted cautiously.

Treatment of localised pustular psoriasis (ACH and PPP) may take the form of topical or systemic therapies, or a combination thereof. While topical agents, such as corticosteroids under occlusion, calcipotriol, fluorouracil or tacrolimus, may initially be successful, patients usually relapse. Systemic treatment is then required. This takes a similar form to the systemic therapies used in GPP, including cyclosporine, acitretin, TNFα blockers and biologics (17,23). However, symptoms are often recalcitrant to these therapies.

While the rarity of ACH is preventing the implementation of controlled studies, the picture is slightly more positive for PPP, where two clinical trials are currently assessing the efficacy of the IL-1 blocker anakinra (EudraCT 2015-003600-23, NCT01794117) and a third has been investigating the IL-17A blocker secukinumab (NCT02008890). Results from the anakinra studies are expected in February 2019 and December 2021 respectively. The secukinumab trial completed in May 2017 but did not meet its primary endpoint and found only weak evidence that the drug could be more effective than a placebo (24). This in turn would suggest a limited role for IL-17A in pathogenesis of PPP.

Finally, recent advances in understanding of the genetic basis of pustular psoriasis have revealed new potential therapeutic targets. Central among these is the IL-36 signalling pathway (25), which is now being targeted by drugs undergoing early phase clinical trials (26,27).

1.1.3 Genetic basis of the disease

Disease-associated alleles have been identified in three genes: *IL36RN*, *AP1S3* and *CARD14*. However, these account for only a minority of cases and so further genetic determinants remain to be identified. While familial aggregation of pustular psoriasis is sometimes seen, disease

presentation can vary widely between individuals of the same family and associations with rare disease alleles of large effect have not been observed.

1.1.3.1 *IL36RN*

IL36RN encodes the IL-36 receptor antagonist, IL-36Ra. This is a soluble molecule which competes with three agonists (IL-36α, IL-36β and IL-36γ) for binding to the IL-1Rrp2 cell-surface receptor. Upon engagement with IL-36α/β/γ, the receptor dimerises with a membrane-bound accessory protein, IL-1RAcP, forming a heterodimer known as IL-36R. This is followed by activation of the NF-κB and mitogen-activated protein (MAP) kinase pathways and expression of a range inflammatory genes, including *IL8*, *IL1*, *IL6* and *TNFα* (28–30). Binding of IL-36Ra to IL1-Rrp2 does not cause dimerization with IL-1RAcP and so signalling via the receptor is prevented.

The IL-36 receptor antagonist and IL-36γ share a β-trefoil structure (12 β-sheets connected by 11 loops) that is characteristic of IL-1 family cytokines (31–33). As there is a high degree of sequence homology between all three IL-36 agonists, it is reasonable to assume that IL-36α and β have similar 3D structures. The regions that differ most between IL-36Ra and IL-36γ are the loops connecting β-sheet 4 to β-sheet 5, β6 to β7 and β11 to β12. Indeed, by exchanging the β4/5 loop in IL-36Ra for that of IL-36γ, the antagonist is converted to an agonist, albeit a weaker one than IL-36γ. Conversely, the long β11/12 loop in IL-36Ra is likely to sterically hinder interaction with IL-1RAcP (31).

All IL-36 cytokines are produced as inactive precursor proteins, requiring N-terminal cleavage by proteases. For example, Cathepsin S cleaves only IL-36γ, but Cathepsin G cleaves IL-36α at Lys3 and IL-36β at Arg5. IL-36Ra is activated by cleavage of the N-terminal methionine by elastase (33,34).

Both IL-1Rrp2 and IL-1RAcP consist of three extracellular immunoglobulin-like domains and an intracellular Toll-interleukin 1 receptor (TIR) domain. The extracellular domains of IL-1Rrp2 form the ligand-binding pocket and also interact with IL-1RAcP to form the IL-36R complex. Detailed in-silico modelling and in-vitro experimental work indicated that within the extracellular region IL-1Rrp2 residues (e.g. Asn-41, Cys-42 and Cys-118) allow for specific recognition of IL-36α, -β or -γ (35).

The NF-κB and MAP kinase pathways are likely to be activated through the recruitment of MyD88 and the Tollip-IRAK complex to the two TIR domains in the receptor complex (36), as this is the mechanism underlying the activation of the IL-1 receptor, which also dimerises with IL-1RAcP. Indeed, MyD88 and Tollip have both been implicated in IL-36 signal transduction (37).

While IL-1Rrp2 and IL-1RAcP will co-precipitate in the absence of IL-36, the addition of IL-36γ more than doubled the level of co-immunoprecipitation. If modified to lack the intracellular TIR domains IL-1Rrp2 and IL-1RAcP are secreted rather than trafficked to the cell membrane. However, they are still able to bind IL-36α/β/γ and the co-receptor and so act as dominant negative regulators of IL-36 signalling (35).

IL36RN variants were first discovered by Onoufriadis et al. and Marrakchi et al., by the analysis of isolated and familial cases of GPP respectively (28,29). These authors also demonstrated that the variants impair IL-36Ra function, thereby upregulating IL-36 mediated signalling (Figure 1.2) and leading to an over-expression of key inflammatory genes (28,29). This correlated well with earlier work in mice, where absence of IL-36Ra in combination with expression of IL-36α in basal keratinocytes led to the development of a severe psoriatic skin phenotype with the addition of intracorneal and intraepithelial pustules (38).

Subsequent studies have identified additional disease-associated alleles. These include missense, stop-gain, splice site and deletion variants, with 20 disease alleles currently listed on

the Infevers database (39). Of note, while the patients described in the Onoufriadis and Marrakchi papers carried homozygous *IL36RN* variants, those same alleles and others have since been seen in patients in the heterozygous state. This raises the possibility that heterozygous variants may be sufficient to cause disease in some individuals, or that in those patients they exist alongside another (yet to be identified) disease allele.

While disease associated *IL36RN* alleles have been seen in all major ethnic groups, the frequency of individual variants varies between populations. In Asians, the most prevalent change is the c.155+6T>G splicing variant, whereas the p.Ser113Leu change is the most common in patients of European descent.

Generalised pustular psoriasis patients who carry one or more *IL36RN* disease alleles collectively account for approximately 30% of cases (40,41). These individuals suffer from a more severe form of the disease, presenting with increased risk of systemic inflammation and earlier age of onset. *IL36RN* variants have been seen less frequently in palmoplantar pustulosis (<5% of cases) than in GPP, which has led some authors to question the association with this form of the disease (41–43). Finally, variants have been seen in cases of ACH, but again the rarity of the condition means estimates of frequency have thus far been difficult to calculate (41,44).

As *IL36RN* screening is increasingly used as a diagnostic test for pustular psoriasis, it is essential that variants identified in patients are correctly classed as disease alleles or benign changes. The ongoing development of drugs to target *IL36RN* deficiency (section 1.1.2) adds to this, as it would be desirable to reliably select patients who would benefit from IL-36 blockade.

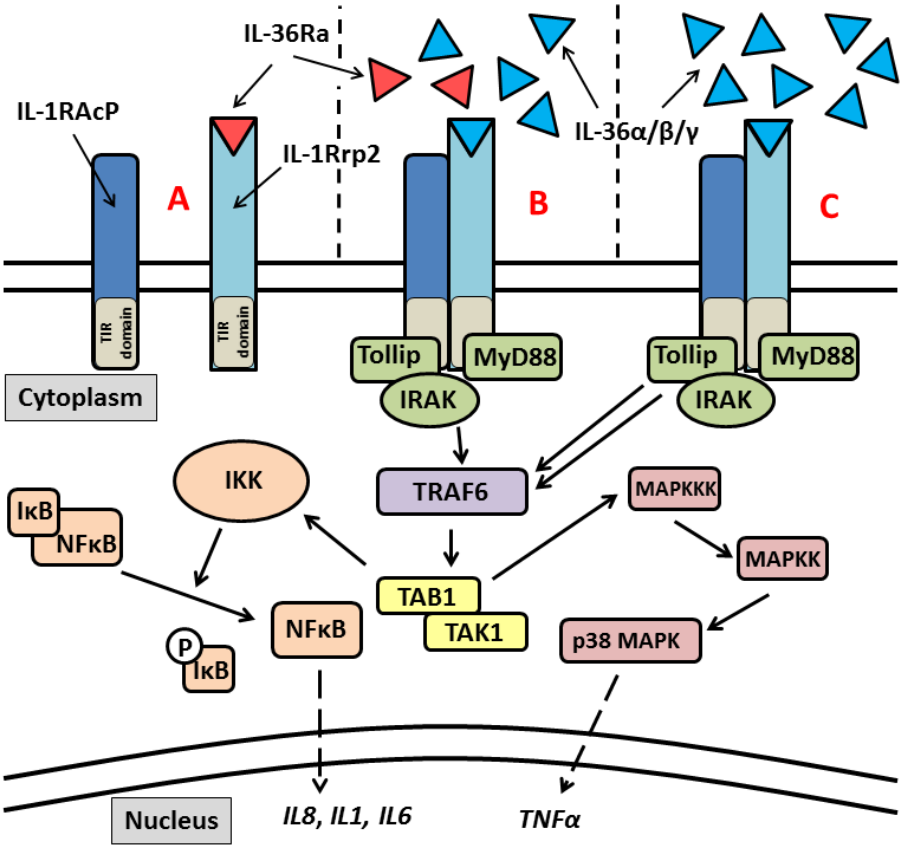


Figure 1.2 Signalling through the IL-36 receptor complex activates NF-κB and MAP kinase pathways

In (A) the receptor antagonist IL-36Ra binds to IL-36R, which remains monomeric and inactive. In (B), IL-36α/β/γ binds to the receptor, which forms a complex with IL-1RAcP. MyD88, Tollip and IRAK are recruited to TIR domains and the NFκB and MAP kinase pathways are activated. This leads to expression of inflammatory genes. In (C), IL-36Ra is absent and so signalling via the IL-36 receptor complex is elevated.

1.1.3.2 *AP1S3*

AP1S3 encodes APσ1C, a small subunit of the AP-1 intracellular trafficking complex (45). Setta-Kaffetzi et al. identified two disease-associated variants, p.Phy4Ser and p.Arg33Trp, through whole-exome sequencing of eight ACH patients and targeted screening in a follow-up cohort (46).

This association has since been confirmed in a larger patient group (47). In addition, three affected individuals have been identified as carrying heterozygous disease alleles in both *IL36RN* and *AP1S3*. Interestingly, one of these patients suffered from a very severe GPP phenotype, with paediatric onset, systemic inflammation and limited drug response. In contrast, a sibling who carried only the *IL36RN* change experienced a much milder form of the disease (41,47). This suggests that *AP1S3* variants may act as a modifier in the presence of *IL36RN* disease alleles.

Variants in *AP1S3* destabilise AP-1 by preventing the interaction between the small and medium-sized subunits of the complex. This disrupts the formation of autophagosomes from the trans-Golgi network, a process in which AP-1 is essential. Impaired autophagy in turn leads to a cytoplasmic accumulation of p62 (Figure 1.3). This is an NF-κB adaptor protein involved in intracellular signalling downstream of toll-like receptor 2/6 and the IL-1 receptor. As a result of p62 accumulation, expression of *IL8, IL1B, TNFA* and *IL36A* is therefore upregulated (47).

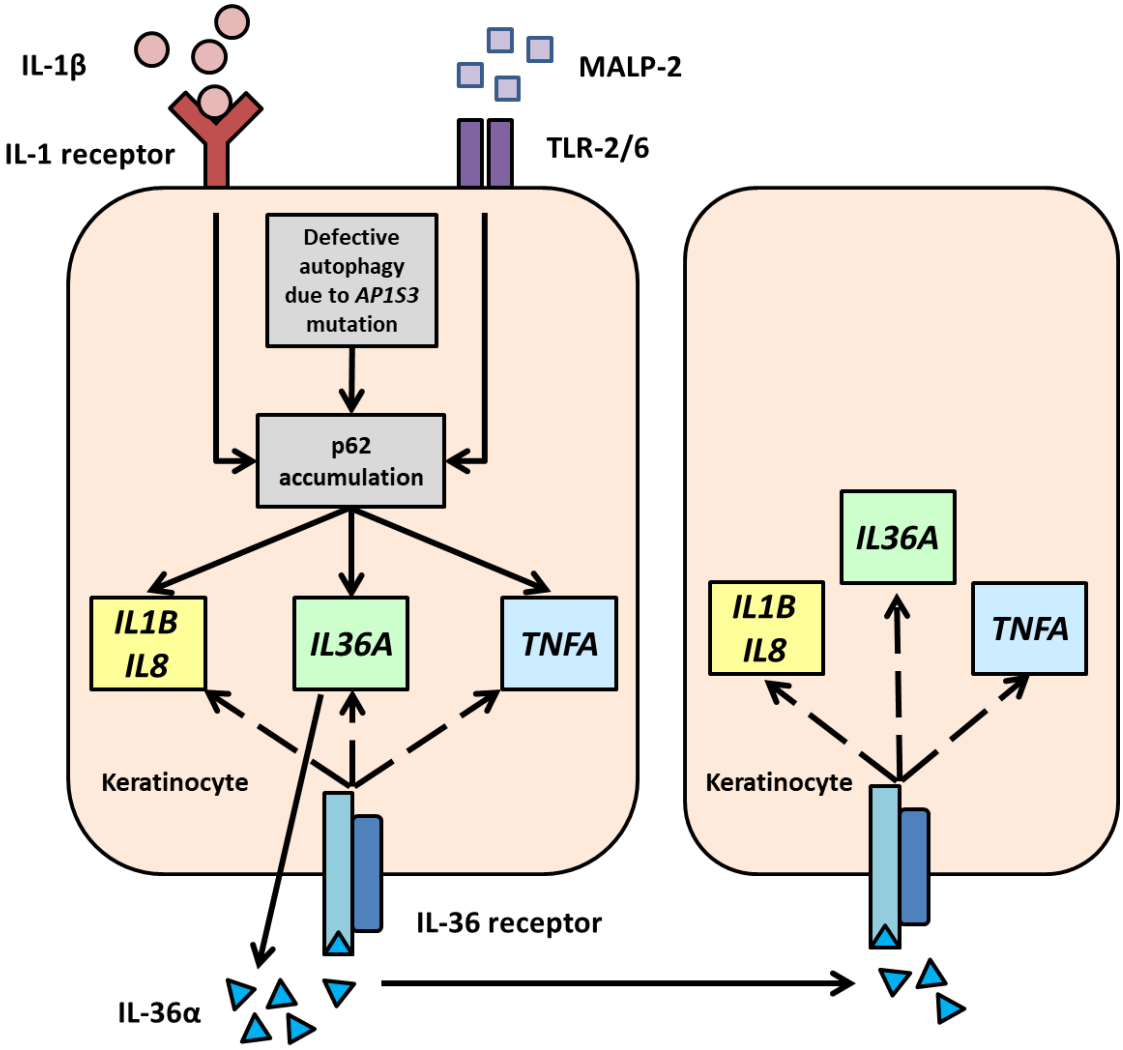


Figure 1.3 The effect of damaging variants in *AP1S3* on production of IL-36α.

AP1S3 disease alleles lead to a disruption of autophagosome formation. This in turn causes accumulation of the NF-κB adaptor protein p62, which results in upregulation of multiple inflammatory genes, including *IL36A*. As a cytokine, IL-36α is able to activate IL-36 receptor-mediated signalling in nearby cells, amplifying the inflammatory response. Figure taken from (47).

1.1.3.3 *CARD14*

Caspase Recruitment Domain Family Member 14 (*CARD14*) is a scaffold protein, expressed in keratinocytes, that is required for BCL10-MALT1 mediated activation of NF-κB signalling. Most *CARD14* variants map to the coiled coil domain, which keeps the protein in a 'closed', inactive, conformation. Disease alleles disrupt this autoinhibitory structure and result in abnormal protein oligomerisation (48), which causes abnormal activation of NF-κB signalling (49) (Figure 1.4).

A missense variant, p.Asp176His, was first identified in Japanese patients with generalised pustular psoriasis and plaque psoriasis (50). Berki et al. subsequently found the same allele in patients of Chinese descent and confirmed the association with pustular psoriasis through a meta-analysis (48). Of note, nearby residues are affected by changes seen in patients with monogenic plaque psoriasis and pityriasis rubra pilaris (PRP) (51,52). Interestingly, multiple studies have demonstrated that some *CARD14* gain of function variants lead to elevated expression of IL-36γ (51,53,54).

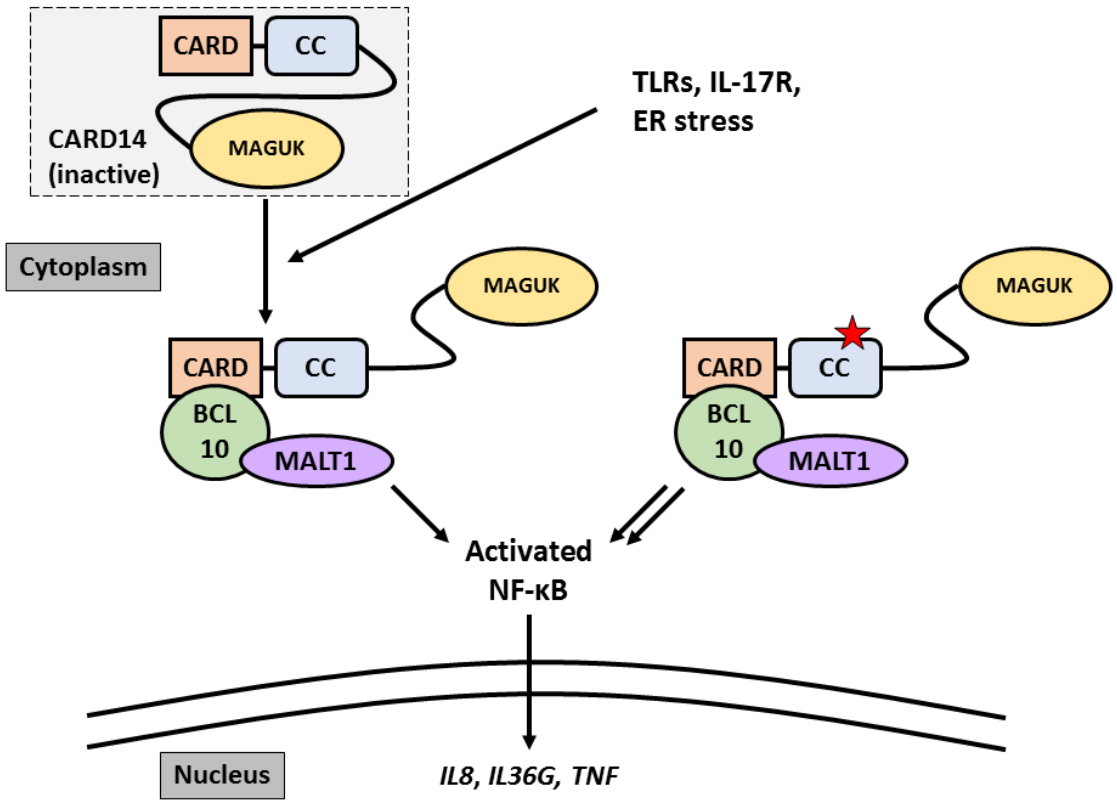


Figure 1.4 An overview of CARD14 signalling

In its inactive form, the 3D conformation of CARD14 prevents the CARD domain from interacting with BCL10. Following signals from toll-like receptors (TLRs), the IL-17 receptor and intracellular organelles such as the endoplasmic reticulum (ER), CARD14 is converted into its active form. The CARD domain can then recruit the BCL10-MALT1 complex, which is able to activate the NF-κB pathway. This leads to expression of pro-inflammatory genes including *IL36G*. Mutated forms of CARD14, indicated here by the red star, are constitutively active, leading to excessive NF-κB activation. Figure adapted from (54,55). CARD: caspase recruitment domain; CC: coiled-coil; MAGUK: membrane-associated guanylate kinase; BCL10: B-cell lymphoma/leukemia 10; MALT1: mucosa-associated lymphoid tissue lymphoma translocation protein 1.

1.1.4 Disease immunopathogenesis

Gene identification studies have clearly demonstrated that IL-36 signalling is a key driver in the pathogenesis of pustular psoriasis; disease alleles are known to affect the receptor antagonist, cause accumulation of IL-36 or result in aberrant activation of a downstream target, the NF-κB pathway. Transcription profiling of patient skin supported this conclusion, highlighting the dominant role of IL-36 cytokines, and more generally the innate immune system, in the development of the disease (25,56).

A significant overlap has been demonstrated between differentially expressed genes (DEGs) in lesional GPP skin and those in keratinocytes stimulated with IL-36α, -β or -γ. (25,56). In addition, genes that are differentially expressed in GPP and upregulated by IL-36 (e.g. *IL8* and *S100A7*) preferentially map to pathways related to infiltration of granulocytes, a characteristic feature of pustular psoriasis (25,56).

A number of mouse models also demonstrate a key role for the IL-36 pathway in psoriasiform inflammation. Firstly, transgenic mice overexpressing *Il1f6* (the murine ortholog of *IL36A*) in basal keratinocytes develop skin lesions with histological features seen in psoriasis, such as hyperkeratosis and leukocyte infiltration of the dermis and epidermis. When the animals are bred with homozygous or heterozygous *Il1f5* (ortholog of *IL36RN*) knockouts, intracorneal and intraepithelial pustules appear within the skin lesions (38).

Secondly, the application of the TLR7 agonist imiquimod on mouse skin causes, over several days, the appearance of psoriasis-like lesions with infiltration of neutrophils and macrophages (57). However, the infiltrate is absent in mice lacking the IL-36 receptor, while it is increased 3-fold in *Il1f5* knockout littermates (58).

Finally, in wild-type mice, pre-treatment with an antibody neutralising the IL-36 receptor significantly reduced the severity of imiquimod-induced psoriasiform lesions (25). Similarly, when human lesional skin was transplanted onto severe combined immunodeficiency (SCID) mice and then treated with an antibody blocking the IL-36 receptor, the psoriasis phenotype improved (59).

Returning to human biology, proteases derived from neutrophils are responsible for the processing of IL-36 cytokines into an active state (34,56). Thus, IL-36/IL-8 mediated neutrophil recruitment leads to an increase in IL-36 activity at the site of inflammation. Interestingly, IL-36 also contributes to the activation of dermal dendritic cells (DCs) and keratinocytes, leading to Th17 cell polarization and chemokine production, respectively (33). Thus, IL-36 cytokines drive powerful feedback loops which propagate inflammation in psoriatic skin (Figure 1.5).

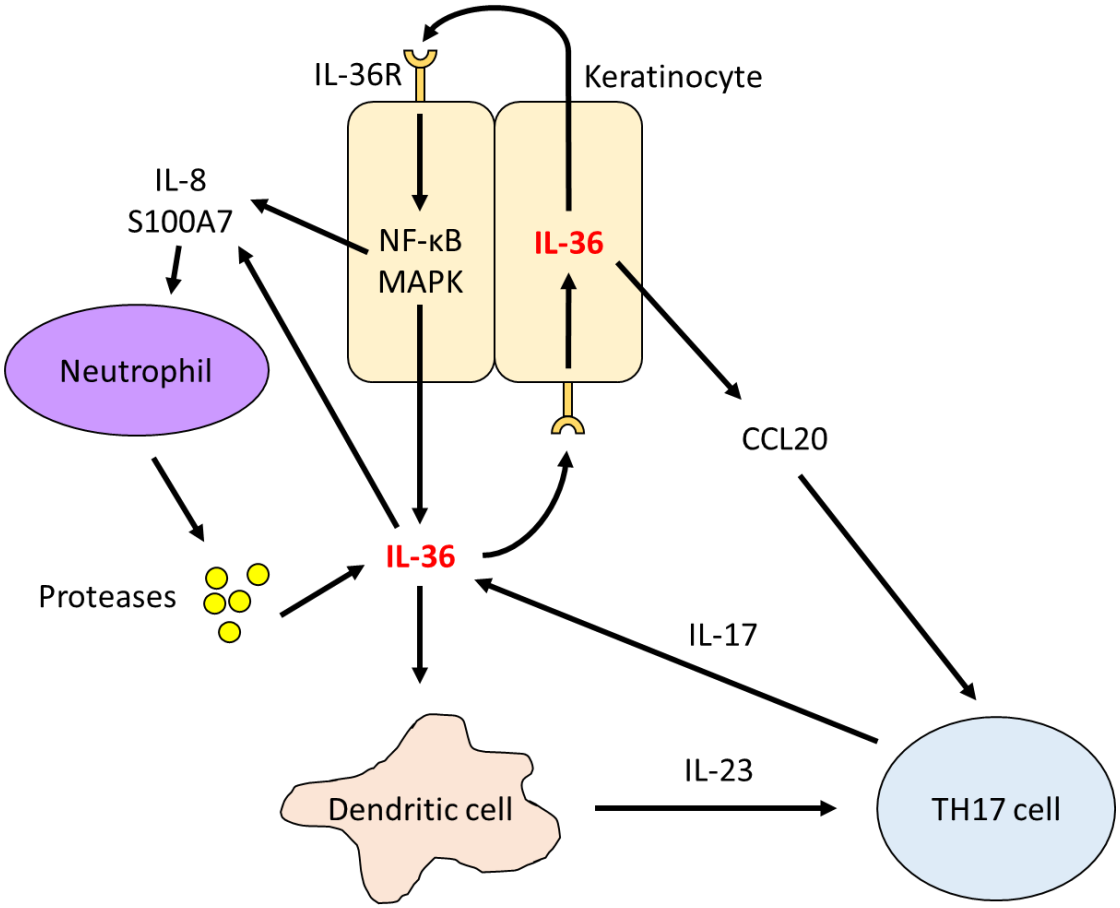


Figure 1.5 The IL-36/Th17 axis in psoriatic inflammation

Figure outlines key cytokines and cell types in development of inflammation driven by IL-36. Keratinocytes (KC) express IL-36, which can upregulate its own expression in a paracrine fashion, but also induce the expression of CCL20, and S100A7 and IL-8. While the former attracts Th17 lymphocytes to the site of inflammation, IL-8 and S100A7 recruit neutrophils. These secrete proteases that convert secreted IL-36 into its active state. Finally, IL-36 activates dendritic cells (DC), leading to IL-23 secretion and polarization of T lymphocytes towards a Th17 phenotype. Figure adapted from (25).

1.2 Gene identification by next generation sequencing

While early advances in gene discovery were made through the use of Sanger sequencing, progress since the late-2000s has been driven by the development of next-generation sequencing (NGS) technology, which simultaneously generates multiple sequence reads from a single DNA molecule. While individual reads may contain errors, the compilation of multiple sequences covering the same region enables NGS platforms to overcome this problem.

1.2.1 Forms of next generation sequencing

The various platforms for NGS can be separated into two categories by the length of the reads they are able to produce. Approaches that produce short reads have historically been cheaper and less error-prone, but long reads are preferable when attempting to sequence repetitive or otherwise complex regions of the genome (60). A summary of the platforms outlined below can be found in Table 1.1.

1.2.1.1 Short read technologies

While there are several different approaches which fall within this category, they all share some similarities in how the DNA is prepared for sequencing. First, the genomic material is fragmented and ligated to adapters, which anchor the DNA to the relevant surface and enable its amplification. Next, large numbers of clonal templates are generated, so that each fragment can be sequenced multiple times simultaneously (60,61).

The most widely used platforms in NGS, to the extent that the market is effectively monopolised, are those produced by Illumina. These employ a ‘sequencing by synthesis’ approach with some similarities to Sanger sequencing, where the fluorophore-labelled deoxynucleotides (dNTPs) are 3’-modified such that a subsequent base cannot be added. In contrast to Sanger sequencing,

however, there are no unmodified dNTPs. In this way, only one new nucleotide can be incorporated into each fragment per cycle. Total internal reflection fluorescence (TIRF) microscopy is used to identify the newly incorporated nucleotide and both the fluorophore and the terminator modification are then removed, a second distinction from classical Sanger sequencing where the modification is not reversible. A new batch of labelled dNTPs are then provided. With Illumina platforms, reads of up to 300bp are possible, with an overall accuracy of >99.5% (60,62).

The Qiagen GeneReader platform uses a very similar approach to Illumina but is mainly applied to the screening of cancer gene panels. DNA fragments are also bound to beads in an emulsion, rather than a solid-phase surface (60).

An alternative version of sequencing by synthesis does not use fluorescently labelled dNTPs. Instead, the IonTorrent platform (ThermoFisher) detects the H^+ ion released upon incorporation of a nucleotide. As this ion is, obviously, the same regardless of the type of nucleotide, each base must be provided individually, removed and replaced by another. If there are multiple adjacent nucleotides of the same type in the DNA fragment, multiple nucleotides will be incorporated in one cycle and the H^+ signal will be stronger (60,63).

In the Roche 454 pyrosequencing platform, the incorporated nucleotide triggered an enzymatic reaction that led to the production of light. These models, however, were discontinued in 2016 (60).

Finally, a different approach to generating short reads is ‘sequencing by ligation’. There are two platforms that use this: SOLiD (Life Technologies) and Complete Genomics. Their key feature is an oligonucleotide probe formed from one (Complete Genomics) or two (SOLiD) known bases attached to an anchor fragment (i.e. a series of universal or degenerate bases). When the probe hybridises with the DNA strand, the anchor fragment is ligated, and the known base(s) are

imaged by detection of fluorophore(s) (64,65). However, these platforms are unpopular, despite high accuracy, due to very short read length (maximum 75-100bp) (60).

1.2.1.2 Long read technologies

These are sometimes referred to as 'third generation' technologies, as they have only come onto the market in the past seven years (66).

The first platform was developed by Pacific Biosciences (PacBio) and utilises a 'single-molecule real-time' (SMRT) approach. This is now capable of producing reads of up to 2kb and while the error rate is high (13%), the DNA molecules can be sequenced multiple times, increasing the overall accuracy. The technique still utilises fluorescently labelled nucleotides, but their incorporation is visualised in real time. In addition, epigenetic modifications can also be detected as they result in an extended interval between incorporations (60,66,67).

Oxford Nanopore Technologies (ONT) have developed sequencers where there is (theoretically) no limit to the length of the DNA molecule that can be sequenced, if the template is of sufficiently high quality. Currently, reads of up to 1Mb have been achieved and so the technology is particularly useful for sequencing long, highly repetitive regions of the genome (66).

The DNA is translocated through a protein nanopore, through which a current is also passed. As the molecule passes through the pore, it disrupts the current in a characteristic pattern. The instrument can distinguish between over 1000 different k-mers and, depending on the type of pore used, detect different epigenetic modifications (66,68).

An issue with nanopore technology is the high error rate (around 15%), but this can be reduced by usage of specialised adaptors that allow the sequencing of both strands in one read (known as the $1D^2$ system). The downside of this is that it doubles the length of time required to sequence the molecule (60,66).

Synthetic long read (SLR) technology is a modification of more traditional short-read approaches that facilitates a more reliable assembly of adjacent fragments. Two companies have released SLR systems: Illumina and 10X Genomics. Briefly, large DNA fragments are partitioned into microtiter wells or an emulsion such that each sample contains very few fragments. Within the partition, the large fragments are sheared and barcoded. After sequencing on a typical short-read platform, the barcodes are used to assemble reads originating from the same large DNA fragment. This method is high throughput and has higher accuracy than true long read approaches, but as it requires a PCR step it suffers from some of the biases (e.g. poorer coverage of GC-rich regions) that affect traditional NGS technologies (60,66).

Table 1.1 Comparison of NGS technologies

Platform[1]	Read length (base pairs)	Throughput	Accuracy	Pros	Cons
Illumina	Up to 300	Up to 900Gb	>99.5%	Compatibility and wide range	Market monopoly, systemic biases
GeneReader (Qiagen)	~100 but driven by gene panels	Dependent on gene panels	~99.5%	Well optimised for cancer gene panels	Limited applications
IonTorrent (ThermoFisher)	200-400	10 – 15Gb	99%	Read length, shorter run-time	High error rate, INDELS problematic
454 pyrosequencing (Roche)	400-700	35 – 700Mb	99%	N/A: discontinued in 2016	
cPal (Complete Genomics)	50-100	8 – 200Gb	99.99%	High accuracy	Short read length
SOLiD (Life Technologies)	75	80 – 320Gb	99.99%	High accuracy	Short read length
PacBio (Pacific Biosciences)	10-15Kb	5 – 10Gb	~99.999% at 20 passes	Detects epigenetic modifications, high accuracy	Expensive, large amount of DNA required

Table continued on following page

Oxford Nanopore Technologies	Dependent on quality of template, current max 1Mb	20Gb-6Tb	~85%, up to ~97% with $1D^2$ system	Detects epigenetic modifications, cost effective	High error rate, large amount of DNA required
TruSeq SLR (Illumina)	~10Kb synthetic length	650Gb – 1.5Tb	>99.5%	Uses pre-existing platforms	PCR-based, same biases as classic short-read seq
10X Genomics	Up to 100Kb synthetic length	800Gb – 1.8Tb	>99.5%	Small amounts of DNA required	PCR-based, same biases as classic short-read seq

[1] Many platforms have multiple variants with varying read length, throughput and accuracy, so the data presented here is a summary. Table contents summarised from (60,66).

1.2.2 Applications of next generation sequencing

Next generation sequencing has numerous applications that are ideally suited to the study of gene regulation. ChIP-seq, for example, is a popular method of investigating DNA-protein interactions. This involves cross-linking DNA to relevant proteins (such as histones), shearing the DNA, capturing protein-bound fragments and then sequencing them (69). A more recently developed technique is the assay for transposase-accessible chromatin (ATAC-seq), which allows genome-wide profiling of chromatin accessibility. A modified transposase cuts the DNA in accessible regions and ligates sequencing adaptors, after which the resulting fragments undergo NGS. The ends of fragments therefore correspond to regions of accessible chromatin (70).

NGS can also be used to profile global gene expression by means of RNA-seq, where RNA is converted to cDNA, sheared and sequenced. The number of reads per gene is normalised relative to reads per kilobase per million (RPKM) (71). By comparing these in treated and untreated cells or in cases vs. controls, it is possible to identify the genes that are differentially expressed between two groups of samples.

However, the most common application of NGS is in the rapid and accurate sequencing of whole genomes or exomes. NGS has therefore massively increased the potential for identification of disease-associated variants, both in monogenic and complex conditions. It not only allows more comprehensive screening of patient DNA but has also enabled population-wide sequencing efforts (e.g. 1000 Genomes Project (72), Exome Aggregation Consortium (73), Genome Aggregation Database (73)).

These have provided researchers with insights into human variation and allowed comparisons between the genomes or exomes of patients and those of healthy individuals of comparable ethnicity. While this can enhance the power of association studies, it is important to bear in

mind that allele frequencies derived from public databases may be biased as a result of population structure. The usage of different platforms and analytical pipelines in cases vs controls can also be a confounding factor, leading to artificially inflated *P* values (74,75).

1.2.3 Processing next generation sequencing data

Whole-exome sequencing (WES) has played an important role in the identification of both *IL36RN* and *AP1S3* as genes that are associated with pustular psoriasis. WES analyses only the coding regions of the genome, which are captured by in-solution hybridisation of RNA probes.

The raw data produced by WES must go through a series of processing and quality control steps before an attempt is made to prioritise candidate disease variants. As the data is generated, it is stored in FASTQ format files, which contain both the sequence of nucleotides and a quality score for each base (75). This is on the Phred scale (1-60) and is calculated as:

$$-10\log_{10}(P(\text{incorrect base call}))$$

A quality control step is then undertaken using tools such as FastQC, which identifies any low-quality bases (generally Phred <20, i.e. *P*(incorrect call) >1%) to be trimmed from the reads, together with the adapters (75–77).

Next, the reads are aligned to a reference genome. Again, there are many available tools, including Bowtie2 and Novoalign. The data is then assessed and processed for a second time, to remove read duplicates (which may be PCR artefacts), improve alignments in regions containing insertions or deletions (INDELs) and recalculate the base quality scores (75).

Finally, a programme such as SAMtools is used to call sequence variants by identifying nucleotides that differ from the reference genome in a sufficient number of reads. These changes are then annotated to known genes (e.g. with ANNOVAR). Additional information such

as the variant zygosity, its population frequency and predicted impact are added at this stage. This data can then be used, alongside prior knowledge of the disease or family history, to prioritise the changes that are most worthy of follow-up (75).

For the purposes of identifying candidate disease alleles, the first step in variant prioritisation is generally to remove synonymous substitutions. Another key consideration is whether the mode of disease inheritance is likely to be dominant or recessive, and therefore whether the variant(s) of interest are likely to be heterozygous or homozygous. If there is sufficient evidence for one or the other, it may be appropriate to restrict the analysis to variants with the relevant zygosity (78).

Next, changes with a minor allele frequency (MAF) that exceed a certain value are removed. While the most appropriate cut-off depends on the prevalence of the disease and the expected penetrance of the mutation, variants with MAF >1-5% are generally regarded as common and excluded from subsequent analysis (78). As described above, data from population-wide WES or WGS is normally used to provide control MAFs. The frequency of rare alleles, however, can vary widely between ethnic groups (79), so it is important that, where possible, MAFs from the appropriate population are used.

Where applicable, subsequent filtering steps may include retaining only variants shared amongst affected family members or with other unrelated patients. A recent publication also demonstrated the value in removing variants commonly seen in private cohorts, even if 'rare' in publicly available datasets, as a method of reducing the number of non-pathogenic variants (80). It may also be useful to consider tissue-specific gene expression patterns or to prioritise variants based on the results of *in-silico* pathogenicity predictions (section 1.3) (75,78). However, such filtering steps must be carried out carefully as they rely on the accuracy of underlying assumptions and the performance of pathogenicity algorithms.

Finally, it is important to check the quality of reads covering the prioritised variants and perform Sanger sequencing of the region, to confirm that the variant base call is not an artefact introduced during the NGS process.

1.2.4 Advantages and limitations of whole exome sequencing

A major advantage of WES is that it is focussed on coding sequences. This makes the data easier and quicker to analyse, with candidate variant lists likely to be shorter. Additionally, there are now a multitude of *in-silico* tools which aim to predict the effects of coding variants (section 1.3.3), allowing for the prioritisation of changes with pathogenic potential. In contrast, the consequences of variants located in non-coding regions covered by WGS can be much harder to decipher. Whole exome sequencing is also cheaper, so a greater number of samples can be sequenced for the same amount of money (60,74).

However, the targeted nature of WES inevitably means that regulatory changes will be missed as these generally lie in non-coding regions. Whole exome sequencing is also inferior compared to WGS when it comes to identifying structural alterations such as copy number variants (CNVs). Reads from a duplicated locus, for example, would likely be mapped to the single copy in the reference sequence. In contrast, WGS data from the same region (particularly if generated on a long-read platform) would contain reads that cover both coding and intergenic regions, allowing detection of the two independent copies of the gene.

Finally, a key variable in the process of WES is the step where the exome DNA is prepared for sequencing. Several hybridisation kits are commercially available, but none are able to capture 100% of the exome. As a result, some target regions will be missing and others will be insufficiently covered (74).

Recently, Cummings et al. demonstrated that combining WES/WGS with RNA-seq of a disease relevant tissue (here skeletal muscle in muscular dystrophy or myopathy) can compensate for some of the limitations of WES and overcome some of the difficulties associated with the interpretation of non-coding changes. For example, they were able to detect an intronic variant within *COL6A1* and demonstrate that it results in abnormal splicing. They also uncovered a substitution in a region of the *NEB* gene that had not been captured during WES sample preparation, and demonstrated that the change disrupted a canonical splice site, causing an extended exon (81).

1.3 Limitations of *in-silico* approaches to pathogenicity prediction

The average human exome diverges from the reference sequence at 20,000 – 23,000 sites. Around half of these differences are non-synonymous, with loss-of function (LOF) changes (INDELs, variants disrupting canonical splice sites or introducing premature stop codons) affecting 250 – 300 genes (72). More stringent analysis has suggested that the typical exome harbours ~100 genuine LOF mutations, resulting in the full inactivation of ~20 genes (82). It is therefore clear that most of the variation affecting the human genome is not disease causing.

This is especially the case for missense single nucleotide variants (SNVs). In fact, the average individual included in the Exome Aggregation Consortium (ExAC) carries 54 alleles which are reported as disease causing in the Human Gene Mutation Database and/or ClinVar. However, 41 of these changes have an allele frequency greater than 1% in at least one ethnic group, suggesting that they are unlikely to be responsible for a rare Mendelian condition (73). As the genotyping of these sites was of high quality, the conclusion was reached that the inflated frequency of 'disease causing' variants was due to classification errors.

1.3.1 Issues raised by variants of unknown significance

Mutation screening may be utilised in clinical settings for the purposes of diagnosis, targeted treatment or genetic counselling. While this can be a rapid and powerful approach if the disease allele(s) have already been validated, interpreting newly identified changes is less straightforward (83,84). Even if a gene (such as *IL36RN*) has been robustly identified as disease-associated, there is a risk that subsequent rare variants will be labelled as disease-causing on the basis of otherwise weak evidence. As NGS-based panels are more frequently used for diagnostic screening and the number of variants seen in a given patient has increased (85), this has become very significant problem.

Similar difficulties present themselves in research settings, where a cohort of patients may undergo whole exome/genome sequencing with a view to identifying disease-associated changes (see section 1.2.2). While successful gene discovery can lead to a better understanding of disease pathogenesis (as has been demonstrated in pustular psoriasis (25,28,46,48)), the analysis of exome profiles is not straightforward. Pathogenicity predictions are key to the filtering of variants, so if the conclusions drawn from *in-silico* tools are inaccurate, the chances of identifying disease genes are severely compromised.

A major question therefore presents itself: as NGS is increasingly used in diagnostic and research settings, and the number of variant alleles inevitably grows, how can the truly pathogenic variants be efficiently and reliably identified?

1.3.2 Classification of variants of unknown significance

In 2015 the American College of Medical Genetics and Genomics (ACMG) and the Association for Molecular Pathology (AMP) published an extensive set of guidelines designed to standardise the clinical interpretation of genetic changes (86). They recommended using a five-tier system whereby variants are classified as pathogenic, likely pathogenic, of uncertain significance, likely benign or benign. Changes are placed into one of these five categories on the basis of 28 different criteria, outlined in Table 1.2 and Table 1.3. The tables also illustrate the range of information which may be accessed when attempting to classify a variant, from the results of functional studies, to those of familial segregation analysis and computational predictions.

The ACMG-AMP guidelines were expanded upon two years later, when a more standardised approach to applying the classifications was proposed and formalised into the InterVar tool (87). This was deemed necessary due to discrepancies in the conclusions reached by different researchers applying the guidelines to the same variants. The discordance in classification is in

part due to the complexity of the guidelines themselves. There is also the issue of personal bias in choosing the databases and *in-silico* tools that will be used to assess a given genetic change (87).

Table 1.2 ACMG-AMP criteria for pathogenic variant classification[1]

Evidence of pathogenicity	Criteria	Details
Very strong	PVS1	Null variant in gene where LOF[2] is known mechanism of disease
Strong	PS1	Same amino acid change as established pathogenic variant, regardless of nucleotide change
	PS2	*De novo* in patient with no family history
	PS3	Well-established *in-vitro*/*in-vivo* functional evidence of damaging effect on gene/gene product
	PS4	Variant is significantly more prevalent in cases than controls (OR[3] >5.0)
Moderate	PM1	Located in mutational hotspot and/or critical functional domain that lacks benign variation
	PM2	Absent from controls in ESP[4], 1000 Genomes or ExAC (or extremely low frequency if recessive)
	PM3	If recessive disorder: detected in *trans* with pathogenic variant
	PM4	Protein length changed by in-frame INDEL/stop loss variant
	PM5	Novel missense change causing amino acid change at residue where different amino acid change is pathogenic
	PM6	Assumed *de novo*, but no parental confirmation
Supporting	PP1	Variant gene known to cause disease cosegregates with disease in multiple affected family members
	PP2	Missense variant in gene where missense variants are common mechanism of disease and with low rate of benign variation
	PP3	Multiple lines of supporting computational evidence
	PP4	Phenotype or family history specific for disease with single genetic aetiology
	PP5	Recently reported as pathogenic by reputable source, but independent evaluation not possible

[1]Adapted from (86); [2]Loss of function; [3]Odds ratio; [4]Exome Sequencing Project

Table 1.3 ACMG-AMP criteria for benign variant classification[1]

Evidence of benign impact	Criteria	Details
Stand-alone	BA1	Frequency >5% in ESP[2], 1000 Genomes or ExAC
Strong	BS1	Frequency greater than expected for disorder
	BS2	Observed in healthy adult individual with relevant zygosity, when full penetrance expected at early age
	BS3	Well-established *in-vitro*/*in-vivo* functional evidence of no effect on gene/gene product
	BS4	Lack of segregation in affected family members
Supporting	BP1	Missense variant in gene where truncating variants known to cause disease
	BP2	Seen with pathogenic variant in *trans* (dominant inheritance) or in *cis* (any inheritance)
	BP3	In-frame INDELs in repetitive region with no known function
	BP4	Multiple lines of computational evidence suggest no impact
	BP5	Found in case with alternate molecular basis for disease
	BP6	Recently reported as benign by reputable source, but independent evaluation not possible
	BP7	Synonymous variant with no splicing impact and nucleotide is not highly conserved

[1]Adapted from (86); [2]Exome Sequencing Project

1.3.3 Commonly used pathogenicity prediction tools

The authors of the ACMG-AMP guidelines noted that their recommendations are of limited utility when evaluating alleles identified in research settings, as the criteria assume that the gene in which the variant lies has already been associated with the condition under investigation (86). Where the disease gene is not known, or where the association is established but there is limited scope to generate *in-vitro*/*vivo* data on a new allele, pathogenicity prediction tools provide a relatively easy way to assess variants.

Predicting the effect of variants in splicing regions is fairly simple, because there are known consensus sequences. Popular tools include Spliceman, MaxEntScan and Human Splicing Finder, all of which compare consensus and variant sites (88–90). If the difference between the two exceeds a certain threshold, the change is deemed to disrupt splicing.

In contrast, determining the pathogenicity (or lack thereof) of missense changes is more difficult. A wide range of tools have been designed which attempt to predict the effect of these alleles, but all have weaknesses.

One of the oldest, and still used, pathogenicity prediction tools is SIFT (Sorting Intolerant From Tolerant). This algorithm is based solely on evolutionary conservation, with the rationale that a change in a highly conserved amino acid is more likely to have a damaging effect. For a given protein sequence, a set of functionally related proteins is compiled using PSI-BLAST and the sequences aligned. SIFT then calculates the scaled probability of seeing any of the 20 possible amino acids at each position. Substitutions with a scaled probability value below a threshold (normally 0.05) are classed as damaging (91).

Another tool based on a similar approach is PROVEAN (Protein Variation Effect Analyzer), although this algorithm also takes into account the conservation levels in the sequences flanking

the variant. It is therefore able to assess the effects of INDELs and multiple substitutions, in addition to SNVs. The PROVEAN score is an average of delta alignment scores from pairwise alignments between the variant sequence and the wild type and related sequences. It is then compared to a threshold (normally -2.5), with variants that have a score below this classed as deleterious (92).

PolyPhen-2 is a more advanced tool, created using a naïve Bayes classifier. This is a simple machine learning algorithm based on Bayes theorem:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

This calculates the probability of *A* happening if *B* has already occurred, e.g. the probability of a variant being pathogenic given a set of information about that nucleotide change. The term naïve indicates that the algorithm assumes all conditions (sources of information about the variant) are independent. While this is not the case in reality, the algorithm is still sufficiently powerful to make valid predictions, especially as the input includes a wide range of features, including sequence alignments but also structural information and protein site/region annotation (93).

PolyPhen-2 was created by training the algorithm on a dataset where both the input (data on conservation, protein annotation, etc) and the output (variant classification) was provided. Two versions of the tool are currently available. One (HumDiv) was trained on alleles annotated in UniProt as causing Mendelian diseases and affecting protein function or stability, while the other (HumVar) used all disease-causing mutations from UniProt. Both training sets also included several thousand variants that were assumed to be nondamaging. As a result, the HumDiv model is more appropriate for the analysis of variants in Mendelian diseases, while HumVar should be used when a more complex genetic aetiology is suspected (94).

MutationTaster is also built on a naïve Bayes classifier and integrates information from a range of sources. Unlike SIFT, PROVEAN and PolyPhen-2, it can also be used to assess variants affecting splicing consensus sequences. The tool uses one of three different prediction models, depending on whether the variant is synonymous/intronic, affects a single amino acid or causes more complex changes to the sequence (95).

One of the more complex and commonly used tools is CADD (Combined Annotation-Dependent Depletion). In developing the program, annotations from a diverse range of sources (e.g. conservation measures, location of DNase I hypersensitivity regions, expression levels, results from other pathogenicity prediction tools) were collated for 14.7 million high-frequency human alleles and from the same number of simulated variants. The assumption was that damaging variants would be depleted from the set of common, naturally occurring alleles but present in the simulated dataset. A support vector machine was trained to differentiate between these two variant sets and a C score was then precomputed for all possible single nucleotide variants in the human exome (96).

Clearly, all tools rely on assumptions about the effect of a variant, such as the importance of individual residues within a protein domain, or how a residue change may affect secondary and tertiary structures. Many are also reliant on the accuracy of pre-existing annotation and experimental data and most have a limited ability to predict the scale of the effect caused by a variant and are limited to analysing variants in isolation rather than taking into account the potential effect of multiple changes.

Unsurprisingly, the more basic pathogenicity prediction algorithms have the most limitations. Primarily, those which only examine evolutionary conservation cannot exploit information on the structure or function of the protein region that is affected by the variant. Their reliability is also limited by the availability of evolutionarily related sequences.

However, newer machine-learning tools also have well-established weaknesses. Specifically, they are vulnerable to bias as a result of type 1 or type 2 circularity. The former occurs when a variant being examined by a tool was present in the training set, which may result in overfitting (the tool is well suited to the training dataset but performs poorly on novel variants). Type 2 circularity occurs when a variant is assessed based on the annotated pathogenicity of other variants in the same genes. The problem with this approach is that the distribution of pathogenic variants in online databases is biased, with most damaging substitutions found within a relatively small number of genes. It is also common for all variants within a gene to be annotated with the same status. Therefore, tools will have limited ability to identify damaging variants in an unannotated gene and incorrectly classify variants that fall within a gene with mixed annotation (97,98).

Machine-learning based tools also tend to periodically update their algorithms, which can lead to changes in variant classification. This can cause problems in both research and diagnostic settings and raise some important questions: How often should the classification be rechecked? At what point in the analysis should the results be considered 'fixed'?

1.3.4 Utility of existing pathogenicity prediction tools

As described above, pathogenicity prediction tools are particularly vulnerable to biased usage and interpretation. There are many of these tools freely available online, and each promises advances in accuracy and precision over its predecessors. However, it is those that are easier to operate (e.g. SIFT, PolyPhen-2 and CADD) that tend to dominate in published research (98). Importantly some tools are based on shared assumptions and theoretical frameworks (e.g. both SIFT and PROVEAN predictions are solely based on conservation, data from PolyPhen-2 was used when training the CADD classifier). Thus, their results should not be considered as independent pieces of evidence (86). Nonetheless, computational predictions may be all that is readily

available to assess a newly identified variant and so efforts must be made, where possible, to mitigate some of their inherent biases and weaknesses.

One option is to assess a variant with multiple, genuinely independent tools and build a consensus classification. However, a study by Ghosh et al found that the combination of tools that best differentiates pathogenic alleles from variants of unknown significance (VUS) performs less well in discriminating VUS from benign changes. The same authors also found that the discordance between tools can cause genuinely benign or pathogenic variants to be erroneously classed as VUS (98).

Another approach that has been explored is to tailor thresholds for pathogenicity within individual algorithms to the gene being examined. In fact, a study of the Human Gene Mutation Database (HGMD) demonstrated very limited overlap between CADD scores associated with disease alleles in different genes (99). This suggests that using a fixed cut-off to determine pathogenicity will result in increased rates of false-positives or false-negatives, depending on the gene in question. The difficulty with designing gene-specific thresholds is that a set of true mutations and benign variants in the gene must already be known.

Ultimately, however, the conclusions that can be drawn from pathogenicity prediction tools alone are limited. The addition of even simple functional assays allows for far more reliable variant classification. Guidugli et al used this approach to assess the clinical relevance of missense variants in the DNA binding domain (DBD) of *BRCA2*, combining a homology-directed DNA repair (HDR) assay with *in-silico* predictions. This enabled them to build a classifier that could be used to predict the effect of new missense variants affecting the DBD (100). In another study, a *PPARG* variant classifier was built after assessing the effects of all possible amino acid substitutions (101), using a high-throughput assay.

It is therefore clear that the integration of functional and computational methods is the best way to approach variant classification.

1.4 Aims

Pustular psoriasis is a rare disease. As a result, patient cohorts and the conclusions drawn from them are limited to some degree by size. Moreover, while *IL36RN* screening is increasingly used as a diagnostic test, its clinical utility is limited by the genetic heterogeneity of the disease and the difficulty of predicting the impact of newly identified sequence variants.

In this context, the study pursued three key objectives, with a view to improving our understanding of pustular psoriasis and facilitating its genetic diagnosis:

1. To identify phenotypic and genetic features that can be used for the stratification of patient cohorts. This was achieved through the systematic analysis of clinical and genetic data collated from a large study resource.

2. To characterise functionally all published *IL36RN* missense changes, using *in-silico* and *in-vitro* methods. A selection of rare substitutions of uncertain significance was also included in this analysis, to determine whether the results obtained from known disease alleles could be used to predict the status of as-yet unclassified variants.

3. To identify novel candidate genes for the disease, by analysing whole exome data from patients with a rare phenotype (acrodermatitis continua of Hallopeau) or unusual disease history (paediatric onset, parental relatedness).

2 Materials and Methods

2.1 Materials

Reagent	Manufacturer	Catalogue number
100bp DNA Ladder	New England BioLabs	N3231S
10X DreamTaq buffer	Thermo Scientific	EP0703
Acrylamide mix	Geneflow Limited	EC-890
Agarose	Severn Biotech	30-10-60
Amersham ECL Western Blotting System	GE Healthcare	RPN2232
Ammonium acetate	Sigma-Aldrich	A2706
Ammonium persulfate	Sigma-Aldrich	A3678
Ampicillin	Sigma-Aldrich	A0166
Autoradiography film	Fujifilm	12715325
Bacto-Agar	BD Diagnostics	214010
Bacto-Tryptone	BD Diagnostics	211705
Bacto-Yeast	BD Diagnostics	212750
BigDye Terminator Sequencing buffer	Applied Biosystems	4336697
BigDye Terminator v3.1	Applied Biosystems	4337454
Bromophenol blue	Sigma-Aldrich	B5525
c-Myc tagged Protein Mild Purification Kit	MBL International/Caltag Medsystems	3305
cOmplete™, Mini, EDTA-free Protease Inhibitor Cocktail	Sigma-Aldrich	11836170001
DMSO	Sigma-Aldrich	D8418
dNTP nucleotides	Fisher Scientific	10520651
DreamTaq DNA polymerase	Thermo Scientific	EP0703
Dulbecco's Modified Eagle Medium (DMEM), high glucose, GlutaMAX™ Supplement	Gibco	61965026

EDTA	VWR/BDH	
Ethanol	VWR	20821.330
Ethidium bromide	Sigma-Aldrich	E1510
Foetal Bovine Serum (FBS)	Gibco	10500064
GeneJET RNA Purification Kit	Thermo Scientific	K0731
Glycerol	Sigma-Aldrich	G7893
Guanidinium hydrochloride	Sigma-Aldrich	G7294
Hi-Di™ Formamide	Applied Biosystems	4311320
HiSpeed Plasmid Maxi Kit	QIAGEN	12663
HiSpeed Plasmid Midi Kit	QIAGEN	12643
Human B2M (Beta-2-Microglobulin) Endogenous Control (VIC™/MGB™ probe, primer limited)	Applied Biosystems	4326319E
Human IL-8/CXCL8 DuoSet ELISA kit	R&D Systems	DY208
Human PPIA (Cyclophilin A) Endogenous Control (VIC™/MGB™ probe, primer limited)	Applied Biosystems	4326316E
IGEPAL	Sigma-Aldrich	I3021
illustra ExoProStar 1-Step	GE Healthcare/Fisher Scientific	11961411
Isopropanol		
Kanamycin	Sigma-Aldrich	K1377
KAPA SYBR qPCR Master Mix	Sigma-Aldrich	KK4604
Laemli dye	0.375 M Tris-HCl (pH 6.8) 9% (v/v) SDS 50% (v/v) Glycerol 5% (w/v) β-mercaptoethanol 0.03% Bromophenol Blue	n/a
LB agar	1% (w/v) NaCl 1% (w/v) Bacto-Tryptone 0.5% (w/v) Bacto-Yeast 1.5% (w/v) Bacto-Agar	n/a
LB medium	1% (w/v) NaCl	n/a

	1% (w/v) Bacto-Tryptone 0.5% (w/v) Bacto-Yeast	
Lipofectamine™ 2000 Transfection Reagent	Invitrogen	11668027
Loading buffer	3% Glycerol 2.5% (w/v) Bromophenol blue ddH_20	n/a
Monoclonal Anti-β-actin antibody	Cell Signalling Technologies	4967S
Monoclonal Anti-c-Myc antibody	Santa-Cruz	9E10
PVDF transfer membrane	Roche	03010040001
Non-denaturing lysis buffer	50mM Tris-HCL (pH7.5) 150mM NaCl 5mM EDTA 1% IGEPAL	n/a
PBS	Gibco	10010015
pcDNA3 vector	Professor Asma Smahi	n/a
pCMV6 entry vector	Origene	n/a
Penicillin-Streptomycin	Gibco	15140122
Polyclonal Anti-Mouse antibody	Agilent/Dako	P0447
Polyclonal Anti-Rabbit antibody	GE Healthcare	NA934V
Ponseau S red dye	Sigma-Aldrich	P3504
Precipitation solution	95% Ethanol 0.12M NaOAc (pH 4.6)	n/a
Precision nanoScript2 Reverse Transcription kit	PrimerDesign	n/a
PrecisionPlus SYBR and ROX qPCR Mix	PrimerDesign	n/a
Proteinase K	Sigma-Aldrich	P2308
QIAPrep Spin Mini Kit	QIAGEN	27104
QuikChange Lightening Site-Directed Mutagenesis Kit	Agilent	210518
Recombinant Human IL-36 alpha/IL-1F6 (aa 6-158) Protein	R&D Systems	6995-IL-010

RNaseZap™ RNase Decontamination Solution	Invitrogen	AM9780
RPMI 1640 Medium	Gibco	21875034
Running buffer	250mM Trizma-Base 1.92M Glycine 1% SDS	n/a
SDS	Sigma-Aldrich	L4509
Skimmed milk powder	Tesco	n/a
Sodium acetate	Sigma-Aldrich	S8625
Sodium chloride (NaCl)	VWR chemicals	27810.295
TaqMan™ Universal Master Mix II	Applied Biosystems	4440040
Tris-buffered saline (TBS)	1.5 M NaCl 0.2 M Trizma-base	n/a
TEMED	Sigma-Aldrich	T9281
Transfer buffer	250mM Trizma-Base 1.92M Glycine	n/a
Tris-EDTA (TE)	0.1 mM EDTA 10 mM Tris-HCl	n/a
Tris-HCL	Sigma-Aldrich	T2569
TrypLE™ Express Enzyme (1X)	Gibco	12605010
Tween 20	Sigma-Aldrich	P7949
β2M PrimerDesign probe	PrimerDesign	54684
β-mercaptoethanol	Sigma-Aldrich	M-3148
Trizma-Base	Sigma-Aldrich	T1503
Glycine	Sigma-Aldrich	G7126
10X Tris-Borate-EDTA (TBE)	Geneflow Limited	B9-0020
Western blotting filter paper	BioRad	1703932

2.2 Study resource

2.2.1 Ethical approval

This study was carried out in accordance with the Declaration of Helsinki. Patients were recruited as part of the FAP (Functional Annotation of Psoriasis susceptibility alleles) study, which was approved by the London - Chelsea Research Ethics Committee (reference 14/LO/2169, 20th January 2015) or the PLUM (Pustular psoriasis: eLucidating Underlying Mechanisms) study, which was approved by the London - London Bridge Research Ethics Committee (reference 16/LO/2190, 30th January 2017). A small subset of PPP patients were ascertained through the APRICOT (Anakinra for Pustular psoriasis: Response in a COntrolled Trial) clinical trial, which was granted ethical approval on 1st April 2016 and assigned EudraCT number 2015-003600-23. The ethics committees of collaborating institutions also granted the required approvals for this study. All patients provided informed written consent for the use of their DNA and clinical data.

2.2.2 Patient ascertainment

This study examined a total of 947 cases of pustular psoriasis, including 6 families, each with two affected individuals. Table 2.1 displays summary information on the isolated cases and further information on familial cases is provided in Table 2.2. Differing subsets of patients were utilised throughout in the study; the demographics of these groups are outlined alongside the relevant results.

All patients were assessed by trained dermatologists and clinical information was recorded in a standardised case report form, included in Appendix 1. Pustular psoriasis was diagnosed and classified based on clinical examination, with the criteria set out by the European Rare And

Severe Psoriasis Expert Network (ERASPEN) (1) used in 581 cases. Subject ethnicity was self-declared.

2.2.2.1 British and Irish pustular psoriasis patients

A total of 329 patients (39 GPP, 7 ACH, 264 PPP and 19 with multiple forms of the disease) were recruited through British or Irish centres (St John's Institute of Dermatology, London; University of Manchester; Glasgow Western Infirmary; St Mary's Hospital, Portsmouth; Our Lady's Children's Hospital, Dublin; St Luke's Hospital, Bradford; University Hospital of North Durham; Royal Victoria Infirmary, Newcastle; Basildon Hospital; University Hospital of Wales, Cardiff; Poole Hospital; Worthing Hospital; Russell's Hall Hospital, Dudley; The Royal Shrewsbury Hospital; Whipps Cross University Hospital; Birmingham Children's Hospital; University Hospital, Galway; Leicester Royal Infirmary; Peterborough City Hospital; Ninewells Hospital, Dundee; Broadgreen Hospital, Liverpool; Bristol Royal Infirmary). DNA was available from 325 of these patients.

2.2.2.2 European Rare And Severe Psoriasis Expert Network patients

Clinical information for 416 patients (58 GPP, 17 ACH, 338 PPP, 3 with multiple disease forms) was provided by recruiting centres affiliated with the European Rare And Severe Psoriasis Expert Network (Medical University of Vienna; Tartu University Hospital; Beni Suef University; University Hospital Zurich; University Medical Center, Schleswig-Holstein, Kiel; Hospital Germans Trias I Pujol, Barcelona; Hospital Universitario Virgen Macarena, Seville; Complejo Asistencial Universitario de León; Hospital 12 de Octubre, Madrid; Universitari Vall d'Hebron, Barcelona). For a subset of these (63 patients) we also received DNA.

2.2.2.3 Other European recruiting centres

An additional 39 patients (11 GPP, 3 ACH, 25 PPP) were ascertained independently of ERASPEN by multiple centres across Europe (University of Szeged; European Registry of Severe Cutaneous Adverse Reactions; Helsinki University Central Hospital; IMAGINE Institute, Necker-Enfants Malades Hospital, Paris; Radboud University Medical Centre, Nijmegen; Universitari Vall d'Hebron, Barcelona; Geneva University Hospital). All provided DNA samples.

2.2.2.4 Non-European recruiting centres

Outside of Europe, 163 patients (156 GPP, 1 ACH, 6 with multiple disease forms) were ascertained at multiple North American, Australian and Asian institutions (Hospital Sultanah Aminah, Johor Bahru, Malaysia; National Skin Centre, Singapore; Royal Prince Alfred Hospital, Camperdown, Australia; Hacettepe University, Turkey; University of Washington, Seattle, USA; National Institutes of Health, Bethesda, USA). All provided blood for DNA isolation. Five GPP kindreds, comprising 10 individuals in total, were identified at the Hospital Sultanah Aminah, Johor Bahru, Malaysia. An additional pair of siblings affected by GPP were identified in Australia. Details of familial cases can be found in Table 2.2.

Table 2.1 Summary of isolated pustular psoriasis cases

Disease form	No. of cases	Sex	Concomitant plaque psoriasis	Age of onset	Ethnicity
Generalised pustular psoriasis	264	169 Female 91 Male 4 Unknown	138/251 (55.0%)	Mean: 31.3yrs Min: 1mth Max: 87yrs	African (n = 53); East Asian (n = 45); European (n = 48); Malay (n = 88); South Asian (n = 21); Other (n = 9)
Acrodermatitis continua of Hallopeau	28	17 Female 11 Male	12/28 (42.9%)	Mean: 51.8yrs Min: 2yrs Max: 83yrs	African (n = 3); East Asian (n = 1); European (n = 22); Other (n = 2)
Palmoplantar pustulosis	627	475 Female 134 Male 18 Unknown	86/546 (15.8%)	Mean: 43.8yrs Min: 4yrs Max: 81yrs	African (n = 23); East Asian (n = 3); European (n = 558); South Asian (n = 8); Other (n = 35)
Multiple forms of pustular psoriasis	28	17 Female 10 Male 1 Unknown	12/28 (42.9%)	Mean[1]: 31.8yrs Min: 2yrs Max: 61yrs	East Asian (n = 4); European (n = 20); Malay (n = 1); South Asian (n = 2); Other (n = 1)

mth: month; PV: psoriasis vulgaris; yrs: years; [1] Age data based on patient age at the onset of first form of pustular psoriasis

Table 2.2 Summary of familial GPP cases

Patient ID	Relationship	Ethnicity	Sex	Age of onset (yrs)	Plaque psoriasis
8GPP1	Sisters	Chinese	Female	32	Y
8GPP2			Female	19	N
16GPP1	Aunt and niece	Chinese	Female	25	Y
16GPP2			Female	17	Y
23GPP1	First cousins once removed	Malay	Female	52	Y
23GPP2			Female	28	Y
38GPP1	Father and daughter	Malay	Male	72	Y
38GPP2			Female	30	Y
41GPP1	Sisters	Malay	Female	52	Y
41GPP2			Female	53	N
OVS0015	Sisters	European	Female	0.9	N
OVS0014			Female	34	N

2.2.3 Control populations

Control populations were utilised at several stages throughout this study. For initial whole exome data filtering, global allele frequencies derived from the Exome Aggregation Consortium (ExAC) database (n=61,406) were used. For gene association tests, allele frequencies from relevant ethnic groups were required. The populations represented were: British (3,781 individuals), Non-Finnish Europeans (63,369 individuals), Malays (96 individuals), South Asians (15,391 individuals) and East Asians (9,435 individuals). Details of control populations are outlined in Table 2.3.

2.2.4 DNA extraction and storage

DNA was extracted from 10ml whole blood by technical employees at St John's Institute of Dermatology. Briefly, ice cold distilled water was added to the sample to lyse red blood cells, and the sample tubes were spun at 1250g for 20min at 4°C. The supernatant was removed and the pellet resuspended in 0.1% IGEPAL. After a second spin under the same conditions, the pellet was resuspended in 6M guanidinium hydrochloride and 7.5M ammonium acetate. The solution was vortexed, supplemented with 10mg/ml proteinase K and vortexed again. After incubation at 60°C for 90min, 96% ethanol was added, and the DNA spooled out on a Pasteur pipette tip. The DNA was immediately placed in 1X TE buffer and allowed to dissolve overnight. Finally, the DNA was reprecipitated with 3M sodium acetate (pH 5.5) and 96% ethanol before being spooled out as before, allowed to dry and dissolved again in 1X TE buffer. Duplicate aliquots of isolated DNA were stored in 1X TE buffer at -20°C. Prior to storage the concentration of DNA was measured on a NanoDrop ND-1000 Spectrophotometer. Working solutions (25ng/µl) were diluted in ddH_2O and stored at -20°C in 96-well plates.

Table 2.3 Sources of control populations

Ethnicity	Dataset	Reference
Multiple ethnicities	Exome Aggregation Consortium (n=60,706) Exomes sequenced in-house (n=674)	Lek, et al, 2016 (73)
British	UK10K (n=3,781)	UK10K Consortium, 2015 (102)
Non-Finnish European	Exome Aggregation Consortium: Non-Finnish Europeans (n = 33,370) Genome Aggregation Database: Non-Finnish Europeans (n = 63,369)	Lek, et al, 2016 (73)
Malay	Singapore Sequencing Malay Project (n = 96)	Wong, et al, 2013 (103)
South Asian	Genome Aggregation Database: South Asians (n = 15,391)	Lek, et al, 2016 (73)
East Asian	Genome Aggregation Database: East Asians (n = 9,435)	Lek, et al, 2016 (73)

2.3 Genotyping by Sanger sequencing

2.3.1 Polymerase chain reaction (PCR)

2.3.1.1 Primer design

Primers were designed using Primer3 (104,105) with template sequences taken from the Ensembl genome browser (106). The possible location of primers within the template sequence was restricted such that they would amplify both the relevant exon and exon/intron boundaries. Where possible, overlap with single nucleotide polymorphisms was avoided by comparing potential primer sequences to the target region on Ensembl. Oligonucleotides were also examined with Primer-BLAST (107) to ensure there were no off-target binding sites.

In the case of *ZNF33A*, primers for exon 5 (6 pairs) were designed manually due to very high levels of coding sequence similarity between *ZNF33A* and *ZNF33B*. The two sequences were aligned using T-coffee (108) and primers then designed in regions where the most 3' bases were specific to *ZNF33A*. These primer pairs also fulfilled the following conditions, where possible: primer length: 18-22nt; Tm: 50-65°C; GC%: 40-60; product length: <500bp. All primer sequences are listed in Appendix 2.

Oligonucleotides were purchased from Eurofins Genomics or Sigma-Aldrich and upon arrival were reconstituted in ddH_2O to a concentration of 100µM. For use in PCR and Sanger sequencing, primers were diluted to a working concentration of 10µM. Both stock and working solutions were stored at -20°C.

2.3.1.2 PCR conditions

Reactions were carried out with 25ng DNA in a final volume of 15µl, using reagents as listed in Table 2.4. Optimal primer annealing temperatures (Tm) were determined by gradient PCR and are listed in Appendix 2. Cycling conditions are shown in Table 2.5.

.

Table 2.4 Components of PCR reaction mix

Reagent	**Volume (µl)**
ddH_2O	9.85
10X Polymerase buffer	1.5
10µM Forward primer	0.5
10µM Reverse primer	0.5
2µM dNTPs	1.5
5U/µl Taq polymerase	0.15
25ng/µl DNA template	1
Total	**15**

Table 2.5 Cycling conditions for PCR reaction

Step		**Temperature**	**Time**
Initial denaturation		95°C	5min
Denaturation	Repeat 30 cycles	95°C	30sec
Annealing		Tm[1]	30sec
Extension		72°C	30sec
Final extension		72°C	10min

[1]Determined for each primer pair by gradient PCR

2.3.2 Agarose gel electrophoresis

Samples of PCR product (3.5µl) were diluted in 1.5µl 5X loading buffer and loaded into a 2% agarose gel, which also contained 0.5mg/ml ethidium bromide. Hyperladder II was used as a molecular weight marker. After amplification products were separated by electrophoresis, the gel was visualised under UV light and bands checked to confirm success and specificity of the PCR reaction.

2.3.3 Sanger sequencing

A clean-up reaction removed unincorporated primers and nucleotides from a 2µl sample of PCR product. illustra ExoProStar 1-Step (0.125µl diluted in 2.875µl ddH_20) was added to the PCR sample, in a 96 well plate, and the mixture incubated at 37°C for 30min then 80°C for 15min.

The full volume of cleaned PCR product was then used for Sanger sequencing, with addition of the reagents listed in Table 2.6 and under cycling conditions displayed in Table 2.7. In the majority of cases, the primer used for sequencing was one of the pair used for the preceding PCR reaction, but in three cases primers internal to the PCR product were used to obtain cleaner sequences (Appendix 2).

Sequencing products were purified by ethanol precipitation. A precipitation solution (26µl/well, 50ml 95% EtOH + 2ml 3M NaOAc, pH 4.6) was added to each product. The plate was incubated at room temperature for 10min and then centrifuged at 3000rpm for 30min. Supernatants were removed by inversion and pellets washed in 100µl 70% ethanol before being centrifuged at 3000rpm for 10min. The supernatant was removed as before, and the pellet allowed to air dry.

Finally, the pellet was resuspended in 10µl formamide and denatured by heating to 96°C for 2min. Samples were run on a 3730xl DNA Analyzer (Applied Biosystems) and visualised with Sequencher v.4.9 (Gene Codes, section 2.8.3).

Table 2.6 Components of Sanger sequencing reaction

Reagent	**Volume (µl)**
BigDye Terminator sequencing buffer	1.25
10µM Primer	0.25
BigDye Terminator v3.1	0.25
Cleaned PCR product	5
Total	6.75

Table 2.7 Cycling conditions for Sanger sequencing

Step		**Temperature**	**Time**
Denaturation	Repeat 30 cycles	96°C	30sec
Annealing		50°C	15sec
Extension		60°C	1min

2.4 Plasmid DNA manipulation

2.4.1 Plasmids used

Mutagenesis was carried out on a pCMV6 entry vector that contained a wild type *IL36RN* insert (obtained from Origene). In addition, three pcDNA3 vectors, each containing a mutated form of *IL36RN*, were kindly provided by Professor Asma Smahi of the Imagine Institute, Necker-Enfants Malades Hospital, Paris, France.

2.4.2 Expansion and purification

E. coli containing the desired plasmid were grown on LB agar plates or in LB medium, where necessary supplemented with the required antibiotic (Ampicillin, 100µg/ml, or Kanamycin, 50µg/ml). Components of LB medium and agar are listed in section 2.1. Transformed bacteria were stored as glycerol stocks (250µl sterilised 70% glycerol and 750µl bacteria culture) at -80°C.

Suspensions of *E. coli* were spread onto LB agar either to expand from a glycerol stock or after transformation (see below). After overnight incubation at 37°C, single colonies were picked and propagated in LB medium as outlined in Table 2.8. To recover purified plasmid, QIAGEN HiSpeed Plasmid Midi and Maxi Kits, or the QIAPrep Spin Mini Kit, were used as described in the manufacturer's protocol.

Table 2.8 Expansion of *E. coli* colonies

Application	**LB medium**	**Source of *E. coli***	**Incubation**
Mini-prep with QIAPrep Spin Mini Kit	5ml	Colony picked from LB agar plate	12-16hrs
Inoculation of large cultures and/or creation of glycerol stocks	5ml	Colony picked from LB agar plate	8hrs
Midi-prep with QIAGEN HiSpeed Plasmid Midi Kit	50ml	1:500 dilution of 8hr culture	12-16hrs
Maxi-prep with QIAGEN HiSpeed Plasmid Maxi Kit	150ml	1:500 dilution of 8hr culture	12-16hrs

2.4.3 Transformation of *E. coli*

Ultra-competent *E. coli* (XL10-Gold strain) were transformed by heat shock. For each transformation, 20ng plasmid DNA was added to a 50µl aliquot of competent cells and incubated on ice for 30min. Aliquots were then heat-shocked at 42°C for 45sec and returned to ice for a further 5min. After addition of 500µl antibiotic-free LB medium, bacterial cultures were incubated at 37°C for at least 1hr before being pelleted by centrifugation at 9000rpm for 5min. The pellet was resuspended in 100µl LB medium and spread onto an agar plate containing the relevant antibiotic, then incubated at 37°C overnight.

2.4.4 Site-directed mutagenesis

Mutagenesis primers were designed using the QuikChange Primer Design Program and then compared to the relevant plasmid sequence to ensure absence of off-target binding sites. Primers were ordered from Sigma-Aldrich and reconstituted in ddH_2O to a stock concentration of 100µM. All primer sequences are listed in Appendix 2. The QuikChange Lightening Site-Directed Mutagenesis Kit was used according to the manufacturer protocol. Briefly, a plasmid containing a wild type construct was combined with mutagenesis primers and amplification reagents then incubated under prescribed cycling conditions. Next, the parental plasmid was digested with *Dpn*I before the mutated vector was transformed into XL10-Gold ultracompetent cells by heat-shock.

After incubation in antibiotic-free medium at 37°C for 1hr, bacteria were spread onto agar plates containing the relevant antibiotic and incubated at 37°C overnight. Isolated colonies were picked, propagated and purified as described in section 2.4.2. The CMV promoter, cDNA insert and polyadenylation regions of the plasmid were Sanger sequenced (as outlined in section 2.3.3, excluding the initial clean-up reaction) to confirm only the desired mutation was present.

2.5 Cell culture

2.5.1 Cell lines

Two cell lines were used, both grown in a NuAire Air-Jacketed Automatic CO2 Incubator (NU-5500). This provides a humidified atmosphere at 37°C, with 5% CO_2. Cell cultures were handled aseptically under a laminar flow hood.

HEK293 (human embryonic kidney) cells were cultured in Dulbecco's Modified Eagle Medium (DMEM) supplemented with 2mM L-Glutamine, 25mM D-glucose, 10% Foetal Calf Serum (FCS), 50U/ml of penicillin and 50µg/ml of streptomycin. HeLa-IL-36R cells (stably transfected with the IL-36 receptor, kindly provided by Prof Seamus Martin, Trinity College Dublin, Ireland) were cultured in RPMI 1640 Medium, supplemented with 11mM D-glucose, 5% Foetal Calf Serum (FCS), 50U/ml of penicillin and 50µg/ml of streptomycin.

For long-term storage, 1-2x10^6 pelleted cells were suspended in 1ml freezing medium (90% FCS and 10% DMSO, filter sterilised), placed in a Mr. Frosty container for 24hr at -80°C then transferred to liquid nitrogen. When removed from storage, aliquots were thawed for 2min at 37°C then combined with 4ml complete medium. Cells were pelleted by centrifugation at 1500rpm for 5min, the supernatant aspirated and the pellet resuspended in an appropriate volume of complete medium. After 24hrs growth in a T25 or T75 flask (dependent on the number of cells that had been frozen), cells were split into a new flask.

To split cells, the medium from a confluent (80-90%) T75 flask was removed and the cells briefly washed with 5ml PBS. To detach cells, 3ml of TrypLE was added and the flasks incubated either at room temperature for 3min (HEK293) or at 37°C for 5min (HeLa-IL36R). Complete medium (7ml) was added to deactivate TrypLE and the cells pelleted by centrifugation at 1500rpm for

5min. Finally, the supernatant was aspirated, the pellet resuspended in 10ml complete medium and a 1ml aliquot transferred to a new T75 flask with 9ml complete medium. If splitting cells from a T25 flask into another of the same size, the same procedure was followed, but volumes were halved. If moving cells from a T25 flask to a T75 flask, cells were split at a 1:3 dilution.

2.5.2 Transfection

Cells (HEK293 or HeLa-IL36R) were transfected with plasmid DNA using Lipofectamine2000. Cells were seeded at a density of $0.6x10^6$/well, in 3ml complete DMEM, in a 6-well plate. Twenty-four hours after seeding, the medium was removed and replaced with 2ml antibiotic-free DMEM. Lipofectamine2000 was then used according to the manufacturer protocol, by diluting plasmid DNA at a final concentration of 2.5µg/well and Lipofectamine2000 at a final concentration of 5µl/well, both in serum- and antibiotic-free medium. Cells were harvested for protein purification or Western blot 24hrs after transfection (section 2.7.1 and section 2.7.3).

2.5.3 Stimulation with IL-36α

HeLa-IL-36R cells were seeded in a 48-well plate at a density of $0.1x10^6$ cells/well, in 500µl complete RPMI medium. After 24hrs, the medium was removed. For each IL-36Ra construct (wild-type or mutagenised), a 150µl master mix including 600ng purified protein in complete medium was split equally between two wells. The plate was incubated for 30min at 37°C, then 25µl of IL-36α (20ng/ml) was added to each well for a final concentration of 5ng/ml. Elution buffer from the protein purification kit (section 2.7.1) was diluted in complete medium and added to both negative and positive control wells in each plate. The negative control was not stimulated with IL-36α, while the positive control was stimulated as described above. Supernatants were harvested 4hrs after stimulation, spun at 1500rpm for 10min to pellet cell debris and stored at -80°C.

2.6 RNA expression studies

2.6.1 RNA extraction

RNA was extracted using the GeneJET RNA Purification Kit, according to manufacturer's instructions, and eluted in 30µl RNase-free water. Steps involving β-mercaptoethanol were carried out under an extraction hood and all surfaces and equipment were cleaned with 70% ethanol and RNaseZap prior to use. Eluted RNA was quantified on a NanoDrop ND-1000 Spectrophotometer then stored at -80°C.

2.6.2 cDNA synthesis

The Precision nanoScript2 Reverse Transcription kit was used, with oligo-dT primers, to generate cDNA from 500ng total RNA. Primers were annealed to RNA during a 5min incubation at 65°C, then cDNA was synthesised during a 20min incubation at 42°C. The reaction was heat-inactivated at 75°C for 10min and 30µl RNase-free water was added for a final volume of 50µl. cDNA was stored in the short-term at -20°C and for the long-term at -80°C.

2.6.3 Real-time PCR

Multiple kit/probe combinations were used to measure transcript levels by real-time PCR. A TaqMan probe and mastermix were used for the housekeeping gene *PPIA*. A second housekeeping gene, *β2M*, was measured with either a TaqMan probe and mastermix or a PrimerDesign probe and PrecisionPLUS SYBR qPCR Master Mix. Transcripts of interest were measured with the KAPA SYBR FAST Master Mix (*ARFGAP2*, *IL8*) or the PrecisionPlus SYBR qPCR Master Mix (*IL36RN*, *IL1RL2*). The primers amplifying these target genes were designed in-house. The sequences (listed in Appendix 2) were required to span an exon-exon junction and to have

similar melting temperatures within each pair (maximum 3°C difference). Primer specificity was validated by agarose gel electrophoresis, which was also used to exclude the formation of primer dimers.

Reagent volumes for all real-time PCR reactions are shown in Table 2.9, and cycling conditions are listed in Table 2.10. All reactions used 2µl cDNA, were set up in duplicate and run on a 7900HT Fast Real-Time PCR System.

Gene expression relative to the housekeeping gene was calculated using the ΔΔCt method (109). Detection threshold cycle (Ct) values were averaged across technical duplicates and then the ΔCt calculated for each sample [1 in the equations below]. This was then normalised to a selected calibrator sample [2] and, finally, the relative gene expression level (RQ) was calculated [3].

$$\Delta Ct = Ct[sample] - Ct[housekeeping] \quad [1]$$

$$\Delta\Delta Ct = \Delta Ct[sample] - \Delta Ct[calibrator] \quad [2]$$

$$RQ = 2^{-\Delta\Delta Ct} \quad [3]$$

Table 2.9 Components of real-time PCR reaction mixes

Real-time PCR kit	Reagent	Volume (µl)
TaqMan	TaqMan Universal Master Mix II	10
	Probe	1
	cDNA	2
	Nuclease-free water	7
PrecisionPlus (housekeeping gene)	PrecisionPlus SYBR and ROX qPCR Mix	10
	Gene probe	1
	cDNA	2
	Nuclease-free water	7
KAPA SYBR FAST	KAPA SYBR qPCR Master Mix	10
	Forward Primer (1µM)	2.8
	Reverse Primer (1µM)	2.8
	cDNA	2
	Nuclease-free water	2.4
PrecisionPlus (transcript of interest)	PrecisionPlus SYBR and ROX qPCR Mix	10
	Forward Primer (1µM)	2.8
	Reverse Primer (1µM)	2.8
	cDNA	2
	Nuclease-free water	2.4

Table 2.10 Cycling conditions for real-time PCR

Real-time PCR kit	Temperature	Time	
TaqMan/KAPA SYBR FAST	95°C	10min	
	95°C	15sec	x40 cycles
	60°C	1min	
PrecisionPlus	95°C	2min	
	95°C	10sec	x40 cycles
	60°C	1min	

2.7 Protein analysis

2.7.1 Protein purification

Myc-tagged IL-36Ra proteins were isolated with the MBL c-Myc tagged Protein Mild Purification Kit. The manufacturer's protocol was scaled down to take into account the lower number of transfected cells ($3.6x10^6$ instead of the suggested $8.8x10^6$). Briefly, each myc-tagged cDNA construct was transfected in triplicate in HEK293 cells, using the protocol reported in 2.5.2. Twenty-four hours later, the medium was removed, the cells detached in PBS (2ml/well, 5min incubation at room temp) and centrifuged for 5min at 1500rpm. The cell pellet was washed in PBS and re-centrifuged as described above. Cells were resuspended in 500µl non-denaturing lysis buffer (section 2.1), supplemented with protease inhibitor (EDTA-free). After incubation on ice for 30min, the cell lysate was transferred to a spin column and 20µl Anti-c-Myc bead suspension was added. The spin column was incubated for 1hr at 4°C with end-over-end mixing then centrifuged for 10sec at high speed. After three washes with wash concentrate diluted 1:10 in ddH_2O, the purified protein was eluted with 20µl elution peptide solution. A second elution step was carried out, re-using the 20µl eluate. The concentration of the purified protein was measured on a NanoDrop ND-1000 Spectrophotometer, using the Protein A280 mode with user-entered molar extinction coefficient and molecular weight values (calculated using BioMol.net online calculator (110)). Proteins were then used the same day for a stimulation assay (section 2.5.3).

2.7.2 ELISA

The R&D Human IL-8/CXCL8 DuoSet ELISA kit was used to assay IL-8 levels after stimulation with IL-36α (section 2.5.3). Samples were diluted 1:80 in 0.1% BSA and the ELISA carried out according to the manufacturer's instructions. The absorbance of each sample was then measured at

450nm and 540nm on a BMG Labtech FLUOstar Omega microplate reader. The raw data was processed with BMG Labtech MARS software, using a four-parameter fit to create the standard curve.

2.7.3 Protein expression assay

2.7.3.1 Gel preparation

A 15% separating gel was used as the protein of interest was small (17kDa), in addition to a 5% stacking gel. Components of both gels are listed in Table 2.11. The separating gel was poured first, then 1ml isopropanol added on top to remove any bubbles that had formed. After the separating gel had polymerised the isopropanol was removed, and the stacking gel poured on top. A comb was carefully inserted and, once all had set, the gel was stored at 4°C soaked in running buffer.

2.7.3.2 Western blotting

Twenty-four hours after transfection, HeLa cells were detached in TryplE (1ml/well), centrifuged for 5min at 1500rpm, washed in ice-cold PBS and re-centrifuged. The cell pellet was resuspended in non-denaturing lysis buffer (section 2.1) supplemented with protease inhibitor (EDTA-free) and incubated on ice for 30min.

Samples were combined with 6X Laemli dye and denatured at 95-98°C for 5min, then loaded onto the prepared gel. Electrophoresis was carried out for 2hr in 1X running buffer, at 100V. Separated proteins were transferred to a PVDF membrane in a Biorad Mini Trans-Blot Electrophoretic Transfer Cell filled with transfer buffer and run for 2.5hr at 200mA. Transferred proteins were visualised by soaking the membrane in Ponseau S red dye and the membrane then blocked in 5% milk-1X TBS-0.1% Tween 20 for 1hr at room temperature.

The membrane was then incubated overnight at 4°C with the primary antibody (Monoclonal Anti-c-Myc or Monoclonal Anti-β-actin, diluted 1:1000 in milk). After three washes of 5min in 1X TBS-0.1% Tween 20, the membrane was incubated 1hr at room temperature with an HRP-conjugated secondary antibody (Polyclonal Anti-Mouse or Polyclonal Anti-Rabbit, diluted 1:10000 in milk). The membrane was then washed as before.

Finally, to visualise the protein of interest, the membrane was treated with the Amersham ECL Western Blotting System for 5 minutes. Autoradiography films were exposed to the membrane and developed using an automated film developer (KONICA MINOLTA SRX-101A). The developed film was then scanned, and densitometry analysis performed with ImageJ (111,112).

Table 2.11 Components of Western blot gels

Gel	Component	Volume (ml)
Separating (15% gel)	H_2O	2.3
	30% acrylamide mix	5
	1.5M Tris (pH 8.8)	2.5
	10% SDS	0.1
	10% ammonium persulfate	0.1
	TEMED	0.004
	Total	**10**
Stacking (5% gel)	H_2O	1.4
	30% acrylamide mix	0.33
	1.5M Tris (pH 6.8)	0.25
	10% SDS	0.02
	10% ammonium persulfate	0.02
	TEMED	0.002
	Total	**2**

2.8 Whole exome sequencing

2.8.1 Sample selection and preparation

One hundred and thirteen samples were exome sequenced in the Capon lab between 2011 and 2018. Samples selected for exome sequencing originated from well-phenotyped patients that also fulfilled one of the following criteria: early age of onset, familial disease, rare disease subtype (namely acrodermatitis continua of Hallopeau) or inclusion within the PLUM study/APRICOT trial.

Genomic DNA samples were diluted in 1X TE buffer to a final concentration of 30ng/µl, with concentration confirmed using a Qubit 2.0 Fluorometer.

2.8.2 Whole exome sequencing

All steps of the whole exome sequencing procedure were undertaken by technicians of the Guy's and St Thomas Hospital Biomedical Research Centre (BRC) core genomics facility.

Genomic DNA samples were diluted in 1X TE (3µg in 130µl) then sheared by sonication into 150-200bp fragments. These fragments were purified with AMPure XP beads and repaired with the SureSelect XT Library Prep Kit, which generates blunt 5'-phosphorylated ends. Following a second purification step, the 3' ends were polyadenylated and the sample purified a third time. After ligation of paired end adaptors, each sample was amplified with the Herculase II Fusion DNA Polymerase Kit. Amplification products were assessed for quality using an Agilent 2100 Bioanalyzer (pre-2017) or TapeStation 2200 (2017 onwards) and for concentration with a Qubit 2.0 Fluorometer.

Exome capture was carried out with the Agilent SureSelect Human All Exome Kit v.4 (pre-2017) or v.6 (2017 onwards), according to the manufacturer's protocol. Sequencing libraries were hybridised with the capture library for 24hrs at 65°C. The hybridised DNA was then captured by streptavidin-coated magnetic beads and amplified with indexing primers. After purification, the amplified DNA was analysed with the Bioanalyzer High Sensitivity DNA Assay. Finally, samples were pooled for multiplex sequencing and prepared for cluster amplification. Sequencing was carried out on a HiSeq1000 (pre-2017) or HiSeq3000/4000 (2017 onwards).

Raw data was annotated by a pipeline established by Prof Michael Simpson and now managed by the Guy's and St Thomas' Hospital BRC core genomics facility. Novoalign (Novocraft Technologies) was used to align paired-end reads to the reference genome (hg19), before the removal of duplicates and poor-quality reads with MarkDuplicates and SAMtools (113). The efficiency of capture was assessed with BEDtools (114). Variants were called by SAMtools where covered by 5 or more reads, then annotated with Annovar (115).

2.8.3 Analysis of Sanger sequencing data

Sanger sequencing data was visualised with Sequencher v4.9 and aligned to reference sequences obtained from the Ensembl genome browser (release 91). Detected sequence changes were also annotated by querying Ensembl. Data from the Exome Aggregation Consortium (ExAC) was used to determine the allele frequency of variants.

2.8.4 Pathogenicity predictions

Five tools were used to assess the pathogenicity of variants located in coding regions. These were: PROVEAN (92), SIFT (91), PolyPhen-2 (93), MutationTaster (95) and CADD (96). Variants within splicing regions were assessed with MutationTaster, CADD, Spliceman (88) and Human Splicing Finder (90) (which includes data from MaxEntScan (89)). Sequence changes that were

classified as deleterious by at least two (splicing variants) or three (coding variants) of these algorithms were considered as potentially pathogenic (Table 2.12). Stopgain variants were an exception, as they qualified automatically.

When pathogenicity predictions were required from a wider range of tools, or when information on evolutionary conservation (such as the GERP++ score) was needed, the dbNSFP database (116,117) or the ANNOVAR web server (118) were queried.

Table 2.12 Interpretation of *in-silico* pathogenicity prediction tools

Tool	Output	Score[1]
SIFT	Tolerated	0
	Damaging	1
PROVEAN	Neutral	0
	Deleterious	1
PolyPhen-2	Benign	0
	Possibly damaging	0.5
	Probably damaging	1
MutationTaster	Disease causing	0
	Polymorphism	1
CADD	<10	0
	10 – 20	0.5
	>20	1
Spliceman	<50%	0
	50-75	0.5
	>75%	1
Human Splicing Finder[2]	Probably no impact	0
	Potentially affecting splicing	0.5
	Probably affecting splicing	1

[1]A combined score of at least 3 meant a variant was considered pathogenic, except for stop-gain alleles which qualified automatically; [2]Result based on Position Weight Matrices (HSF) and MaxEntScan

2.8.5 Filtering of whole exome data

Annotated whole exome data (section 2.8.2) was imported into Microsoft Excel and analysed by filtering and prioritisation of variants based on a number of criteria. First, synonymous variants and those annotated as having an unknown effect were removed. Subsequent analysis differed according to the cases included.

2.8.5.1 Multiple unrelated cases

When analysing data combined from multiple unrelated ACH patients, variants were split into two groups: heterozygous and homozygous/compound heterozygous. Both groups were then filtered according to the minor allele frequency (MAF) of individual variants. In the heterozygous group, only variants with a MAF <0.01 were retained. A more relaxed threshold of <0.05 was used for the homozygous/compound heterozygous group. The sources for minor allele frequency were the Exome Aggregation Consortium (73), the 1000 Genomes Project (72), the Exome Variant Server (119) and 700 exomes sequenced in-house.

Next, variants found in only one individual were discarded, along with those likely to be artefacts and those in genes that are very commonly seen in public exome datasets (120). The remaining variants were annotated with pathogenicity predictions as described in section 2.8.4 and those classified as damaging by a majority of tools were retained. An RNAseq dataset generated from unstimulated keratinocytes (25) was used to highlight variants within genes with a reads per kilobase per million (RPKM) of at least 7, placing them in the top 75% of genes by expression level.

2.8.5.2 Isolated case with related parents

For the analysis of a paediatric onset case of GPP, it was assumed that the causative variant would be homozygous as the parents of the proband were first cousins. Therefore, only homozygous variants with a MAF <0.05 were retained; the sources of minor allele frequency were the same as listed in the previous section. Next, variants that, in the original unfiltered data, lay within a homozygosity block of at least 2Mb were retained. This increased the likelihood that the remaining homozygous variants were identical by descent.

As described above, pathogenicity prediction tools were used to prioritise genes containing damaging variants and these genes were then filtered on the basis of their expression in keratinocytes.

2.8.6 Visual validation of whole exome reads

Whole exome sequencing reads relating to the variants that had passed the above filtering steps were visualised with the Integrative Genome Viewer (IGV) (121). The position of each variant of interest was assessed for coverage and quality of reads. Variants covered by fewer than 10 reads or that were in a region containing numerous mismatches were labelled as artefacts and excluded from further analyses.

2.9 Statistical analyses

2.9.1 Comparative and association tests

When examining clinical traits in pustular psoriasis, a binomial test was used to compare the sex distribution amongst patients with that observed in the general population. The distribution of the data for age of onset was assessed with the D'Agostino-Pearson normality test in GraphPad Prism 7**.** A Kruskal-Wallis test, followed by Dunn's multiple comparisons test, was used to analyse differences in age of onset between forms of pustular psoriasis.

To assess for association between clinical traits (e.g. plaque psoriasis) and the disease, or genetic variants and the disease, a Fisher's exact test or chi-squared test were used. The choice of test was dependent on sample size and the number of factors being assessed. A Yates correction for continuity was used with chi-squared tests with one degree of frequency. When examining the association between genetic variants and pustular psoriasis, a one-tailed Fisher's exact test was applied. This was justified as rare deleterious variants would be strongly expected to be enriched amongst patients, not reduced.

When combining ethnicity-specific data on association between genetic variants and pustular psoriasis, a meta-analysis was carried out in Comprehensive Meta-Analysis Version 3 (122). A weighted pooled odds ratio and Z score was calculated.

When comparing densitometry measurements and ELISA readouts between mutant and wild type IL-36Ra, Dunnett's multiple comparisons test was carried out in GraphPad Prism 7.

All statistical tests described in this section, unless otherwise stated, were carried out in R (123,124).

2.9.2 Power calculations

The Genetic Power Calculator (125) was used to assess whether components of the study had sufficient power to detect disease-associated variants of a realistic frequency and genotype relative risk. The disease prevalence was set at 0.09% and D-prime as 1, with the required power threshold defined as 80%.

2.9.3 Regression analyses

Genotype-phenotype correlations were assessed using a linear (age of onset) or logistic (sex, psoriasis vulgaris concurrence) regression analysis. The disease subtype was included as a covariate. Analysis of rates of smoking in PPP and GPP was done using a logistic regression with age of onset and nationality as covariates. All regression analyses were carried out in R.

2.9.4 Significance thresholds

In the case of tests addressing clinical data, *P* values of less than 0.05 were considered significant. The same threshold was used for the analysis of ELISA, Western blot and real-time PCR data. When examining the association between sequence variants and the disease, an exome-wide threshold of $P<2.5x10^{-6}$ was used to compensate for the issue of multiple testing. This value is reached by applying a Bonferroni correction to 20000 independent tests (one test per gene in the genome) each with a *P* value of 0.05 (74).

3 Clinical and genetic analysis of an extended patient cohort

The majority of published findings regarding the clinical presentation or genetic basis of pustular psoriasis are based on the analysis of small patient groups. To address the limitations associated with small sample size, the study collated clinical information from 863 affected individuals, generating a dataset that exceeds by almost 3-fold the size of largest cohort reported to date (n = 323 (41)). A representative patient subset (n = 473) was also screened for variants in known disease-associated genes.

3.1 Cohort selection and demographics

A total of 863 patients were examined (Table 3.1). These individuals suffered from generalised pustular psoriasis (GPP, n = 251), acrodermatitis continua of Hallopeau (ACH, n = 28), palmoplantar pustulosis (PPP, n = 560) or multiple forms of pustular psoriasis (n = 24).

Case report forms based on the ERASPEN diagnostic criteria (1) (Appendix I) were used to phenotype 506/863 of the individuals. The remaining patients were included in the study based on the assessment of an experienced dermatologist and the availability of data on at least one of the following: age of onset, sex or plaque psoriasis concurrence. Ethnicity was self-declared.

Clinical ascertainment occurred at 41 locations. When grouped by country of origin, 95% of patients (823/863) fell into one of eight regional cohorts: UK/Ireland, Malaysia, Austria, Estonia, Egypt, Switzerland, Germany or Hungary (Table 3.2). The majority (n = 591, 68.5%) were of European descent, with substantial numbers also originating from Asia (n = 161, 18.7%) and Africa (n = 78, 9.0%). The distribution of pustular psoriasis subtypes was not the same across all

ethnicities: the majority of Asian participants suffered from GPP, while most European patients had PPP (Table 3.1, Table 3.2). However, this was largely due to ascertainment bias (e.g. some clinicians focused on one subtype over the others) rather than genuine demographic differences.

Table 3.1 Summary description of the patient cohort

	Ethnicity				Sex			Clinical diagnosis						
	European	**Asian**	**African**	**Other[1]**	**Female**	**Male**	**Unkn[2]**	**ACH[3]**	**GPP[4]**	**PPP[5]**	**ACH + GPP**	**ACH + PPP**	**GPP + PPP**	**Total**
Total	591	161	78	33	620	233	10	28	251	560	9	4	11	863

[1]Includes unknown ethnicity (n = 19), mixed ethnicity (n = 4), Middle Eastern (n = 4), Finnish (n = 2), Filipino (n = 1), Hispanic (n = 1), Jamaican (n = 1), Romani (n = 1); [2]Unknown; [3]Acrodermatitis continua of Hallopeau; [4]Generalised pustular psoriasis, [5]Palmoplantar pustular psoriasis

Table 3.2 Breakdown of patient cohorts by ethnicity and diagnosis

		Clinical Diagnosis					
Cohort	**Ethnicity**	**ACH**	**GPP**	**PPP**	**Multiple diagnosis**	**DNA available**	**Total**
UK/Ireland (St John's Institute of Dermatology, London)	European	6	31	169	15		
	Asian	-	3	9	-		
	Other	1	2	19	-	221	**255**
Malaysia (Hospital Sultanah Aminah, Johor Bahru)	Asian	-	130	-	6		
	Other	-	2	-	-	137	**138**
Austria (Vienna)	European	-	-	118	-		
	Asian	-	-	1	-		
	Other	-	-	1	-	-	**120**
Egypt (Beni Suef University)	African	1	51	22	-	-	**74**
Switzerland (University Hospital of Zurich)	European	14	3	53	-	61	**70**
Germany (University Medical Centre Schleswig-Holstein, Kiel)	European	2	-	33	2	-	**37**
Hungary (University of Szeged)	European	-	-	25	-	25	**25**
Estonia (Tartu University Hospital)	European	-	-	103	-		
	Other	-	-	1	-	-	**104**
Europe – other	European	-	7	6	1		
	Asian	-	2	-	-		
	Other	3	6	-	-	14	25
Asia – other	Asian	1	9	-	-		
	Other	-	1	-	-	11	11
Australia/USA	European	-	4	-	-	4	4
Total		**33**	**250**	**569**	**23**	**473**	**863**

3.2 Comparisons of clinical features

The study first focussed on three key clinical features: age of onset, sex and plaque psoriasis concurrence. The distribution of these parameters was comparable between different ascertainment centres (Table 3.3), with the exception of the rate of concurrent plaque psoriasis in GPP, which was much higher in patients ascertained in Malaysia compared to the rest of the cohort (77.3% vs 26.4% in non-Malay GPP).

The 24 individuals with multiple forms of pustular psoriasis were excluded from further analysis, as they could not be unambiguously assigned to a single clinical group.

3.2.1 Age of onset

The earliest age of onset recorded within the cohort was just one month, whereas the latest was 87 years. The mean age of onset was 40yrs, with the median slightly higher at 42yrs. A D'Agostino-Pearson test revealed that the data was not normally distributed ($P < 0.0001$) and so non-parametric tests were used in subsequent analysis.

While paediatric (<10yrs) and very late onset (>70yrs) cases were present in all three disease subtypes, a highly significant difference in median age of onset was observed (Figure 3.1, panel A). Symptoms manifested, on average, earlier in GPP (31.0 ± 19.7 years) than PPP (43.7 ± 14.4; $P = 9.3 \times 10^{-19}$) or ACH (51.8 ± 20.4; $P = 1.2 \times 10^{-7}$).

3.2.2 Concurrent plaque psoriasis

It is well known that patients with pustular psoriasis may also develop plaque psoriasis during their lifetime (1,8,21). Indeed, 29.1% (238/817) of individuals in the cohort suffered from this condition, a much higher prevalence than the 2-5% seen in the general population ($P<2.2 \times 10^{-16}$).

However, when patients were separated by disease subtype, plaque psoriasis concurrence rates differed significantly. Specifically, plaque psoriasis had a lower prevalence in PPP (15.8%) than GPP (54.4%; $P < 2.2 \times 10^{-16}$) or ACH (46.2%; $P = 0.0004$) (Figure 3.1, panel B). This difference remained significant ($P = 0.013$) even after removing the 132 Malaysian GPP patients who had a very high rate of plaque psoriasis.

3.2.3 Sex ratio

Of the 853 patients for whom the sex was known, 620 (72.7%) were female. A significant bias in female to male ratios was observed among GPP and PPP cases (1.7:1; $P = 3.95 \times 10^{-6}$ and 3.5:1; $P < 2.2 \times 10^{-16}$ respectively). The difference between the ratios observed in the two conditions was significant ($P = 5.84 \times 10^{-5}$) (Figure 3.1, panel C), highlighting PPP as the disease with the most pronounced sex bias. An unequal sex distribution was also seen amongst ACH cases (1.5:1) but was not statistically significant.

3.2.4 Smoking status

It is well established that a history of smoking is a risk factor in PPP (17,21). Indeed, 80.7% (264/327) of PPP patients for whom data was available here were current or ex-smokers, compared to 29% (27/93) in the GPP group. The data is difficult to analyse outside of the aggregate as samples become comparatively small. Also, individual resources tend to be heavily skewed towards GPP (e.g. Malay dataset) or PPP cases (e.g. Austrian dataset), hindering a more granular comparison. Despite this caveat, the significance of smoking on PPP patient numbers cannot be overlooked in the aggregate. A logistic regression analysis reveals smoking to be significant when analysed alone ($P < 2.2 \times 10^{-16}$), and still the only significant factor ($P < 0.005$; no other factors had $p < 0.05$) when analysed along with age and nationality. Thus, while the drop in significance suggests that other factors contribute to the probabilistic weightings in the

aggregate, smoking remains the most important factor. This is consistent with data published by many others (126–128).

Interestingly, the concurrence of plaque psoriasis was higher among PPP patients who smoked, compared to their non-smoking counterparts (12.4% vs. 1.6%, P = 0.009). This suggests that smoking not only increases the risk of PPP but may also affect other manifestations of the disease. A larger dataset, allowing us to control for potentially confounding variables, would be required before this association can be confirmed.

Table 3.3 Phenotype characteristics of the major cohorts[1]

		Sex			PV status			Age of onset (yrs)	
Cohort		**Female**	**Male**	**Unknown**	**Has PV**	**No PV**	**Unknown**	**Mean**	**SD**
UK/Ireland	GPP	27 (75.0)	9 (25.0%)	-	16 (41.7%)	17 (44.4%)	3 (13.9%)	31.8	26.3
	PPP	151 (76.4%)	44 (22.6%)	2 (1.0%)	47 (23.6%)	119 (59.5%)	31 (16.9%)	42.4	15.9
Malaysia	GPP	90 (68.2%)	41 (31.1%)	1 (0. 7%)	102 (77.3%)	29 (22.0%)	1 (0.7%)	33.3	17.1
Austria	PPP	103 (85.8%)	17 (14.2%)	-	14 (11.7%)	106 (88.3%)	-	42.8	13.0
Egypt	GPP	21 (41.2%)	30 (58.8%)	-	4 (7.8%)	47 (92.5%)	-	27.9	18.4
	PPP	5 (23.8%)	17 (76.2%)	-	1 (5.0%)	21 (95.0%)	-	37.0	14.2
Switzerland	PPP	39 (73.6%)	14 (26.4%)	-	14 (26.4%)	38 (71.7%)	1 (1.9%)	42.4	15.9

Germany	PPP	25 (76.5%)	6 (17.6%)	2 (5.9%)	3 (9.1%)	29 (87.9%)	1 (3.0%)	42.2	16.1
Hungary	PPP	20 (80%)	5 (20%)	-	3 (12%)	22 (88%)	-	49.5	12.4
Estonia	PPP	83 (79.2%)	20 (19.8%)	1 (1.0%)	-	103 (99.0%)	1 (1.0%)	47.0	10.9

[1]Data only shown where the number of patients within a cohort with a given clinical diagnosis was greater than 25

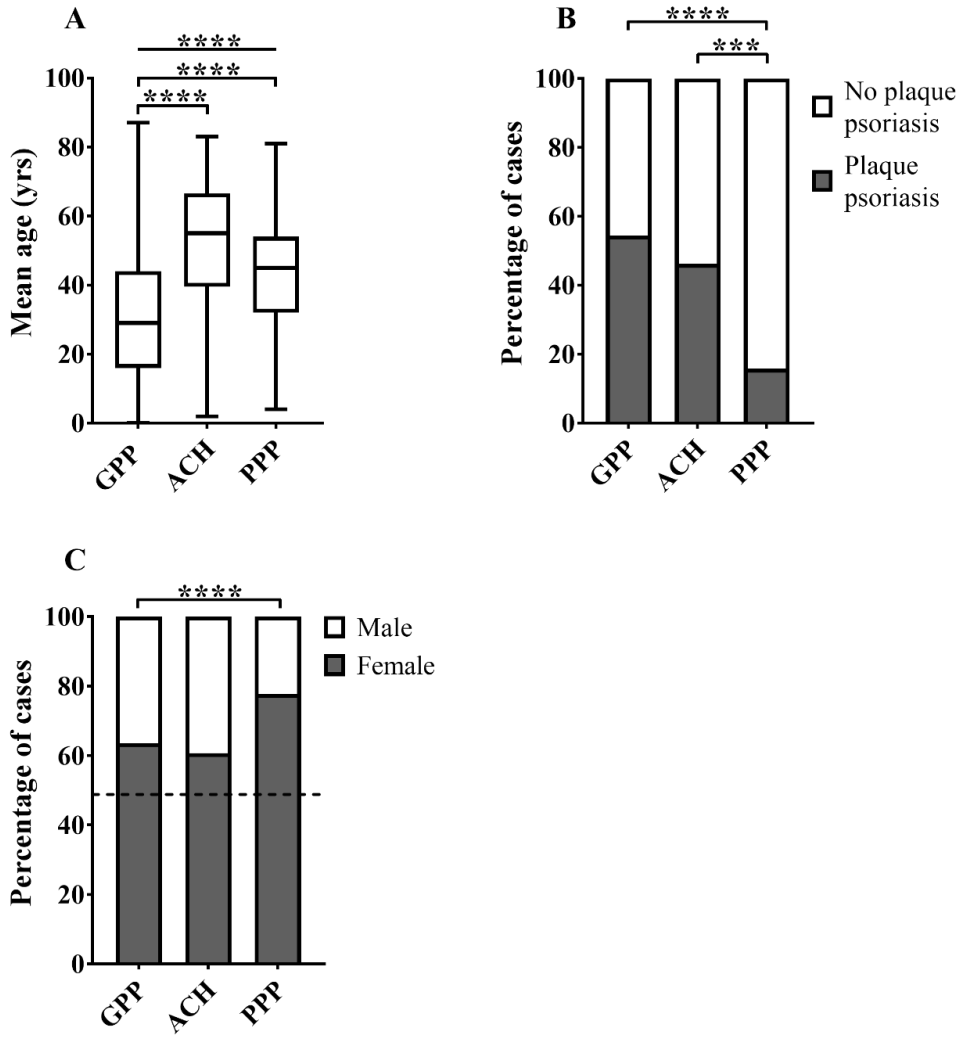


Figure 3.1 Key demographics and clinical features observed in the patient cohort

(A) Disease onset differed significantly across disease subtypes as demonstrated by a Kruskal-Wallis test followed by Dunn's multiple comparison test. (B) Plaque psoriasis prevalence was lowest in PPP. Differences in disease concurrence were analysed with a chi-square test. (C) The ratio of females to males was higher in PPP than in GPP. Differences in the proportion of affected females were assessed using a chi-square test. The dashed line indicates the percentage of females in general population. **P < 0.01; ***P < 0.001; ****P < 0.0001.

3.3 Genetic analysis

The next stage of the study aimed to assess the relationship between key clinical characteristics and patient disease allele status. Within the cohort, DNA was available for 473 of the 863 (54.8%) individuals. Importantly the demographics and clinical features observed in this patient subset were representative of the cohort as a whole (Table 3.4).

3.3.1 Patient screening

The 473 affected individuals included 358 cases who had been previously screened by the Capon group (28,44,46–48) and 115 newly recruited patients. Among the latter, 59 were sequenced in-house and 56 were analysed by members of the Navarini group at University Hospital Zurich, Switzerland. In the majority of cases (397/473), variant screening was carried out by Sanger sequencing of coding regions and exon/intron junctions. While there are newer and more advanced technologies available (e.g. the Nextera XT DNA Library Preparation Kit combined with the MiSeq sequencer), the limited number and small size of candidate genes here allowed for the use of Sanger. This remains one of the most accurate screening protocols. If new candidate genes are discovered, the use of NGS-based techniques will likely prove more advantageous and cost-efficient. The status of the remaining 76 patients was determined by querying whole exome sequencing data previously generated by the Capon group.

3.3.2 Analysis of *IL36RN* variants

The coding region of *IL36RN* extends across four exons and is 468bp in length. Here, *IL36RN* variants were identified in 66/467 (14.1%) patients (45 GPP, 4 ACH, 12 PPP, 5 multiple diagnoses) (Table 3.5), with a combined disease allele frequency of 0.11. While the majority of *IL36RN*-positive patients (36/66, 54.5%) harboured biallelic (homozygous or compound heterozygous)

changes, monoallelic changes were observed in 30/66 individuals. The most common disease alleles were the missense p.Ser113Leu (c.338C>T) variant (n = 42/102) and the c.115+6T>C splicing change (n = 53/102). A previously unreported c.115+5G>A substitution was also identified. This affects a highly conserved residue within the donor splicing motif of intron 3-4 (129) and so is predicted to be deleterious (Figure 3.2). All remaining changes had already been described in the literature.

Table 3.4 Clinical and demographic features of the 473 patients who were screened for variants

		Sex			PV status			Age of onset (yrs)	
Diagnosis	**Cohort**[1]	**Female**	**Male**	**Unknown**	**Has PV**	**No PV**	**Unknown**	**Mean**	**SD**
GPP	Screened cases	128 (67.0%)	59 (30.9%)	4 (2.1%)	125 (65.5%)	56 (29.3%)	10 (5.2%)	31.4	19.6
	Entire dataset	157 (62.5%)	90 (35.9%)	4 (1.6%)	131 (52.2%)	110 (43.8%)	10 (4.0%)	31	19.7
ACH	Screened cases	14 (60.9%)	9 (39.1%)	-	10 (43.5%)	11 (47.8%)	2 (8.7%)	55.1	19.2
	Entire dataset	17 (60.7%)	11 (39.3%)	-	12 (42.9%)	14 (50.0%)	2 (7.1%)	51.8	20.4
PPP	Screened cases	182 (75.5%)	57 (23.7%)	2 (0.8%)	57 (23.7%)	155 (64.3%)	29 (12.0%)	43.2	15.2
	Entire dataset	431 (77.0%)	124 (22.1%)	5 (0.9%)	83 (14.8%)	443 (79.1%)	34 (6.1%)	43.7	14.4

[1] The features observed in the broader dataset are shown for reference

Table 3.5 Variants observed in *IL36RN* positive patients[1]

Sample ID[2]	Diagnosis	Ethnicity	*IL36RN* genotype[3,4] cDNA; protein	Chromosomal location
041KJO86	ACH	European	c.338C>T/-; p.Ser113Leu/-	chr2:g.113062547/-
GBR0019	ACH	European	c.304C>T/c.338C>T; p.Arg102Trp/p.Ser113Leu	chr2:g.113062513/chr2:g.113062547
OVS0026	ACH	East Asian	c.338C>T/c.338C>T; p.Ser113Leu/p.Ser113Leu	chr2:g.113062547/chr2:g.113062547
IL14	ACH	European	c.115+6T>C/c.[115+6T>C;227C>T]; p.Arg10Argfs*1/p.[Arg10Argfs*1;Pro76Leu]	chr2:g.113060943/ chr2:g.[113060943;113062235]
017KSA97	GPP	European	c.338C>T/c.338C>T; p.Ser113Leu/p.Ser113Leu	chr2:g.113062547/chr2:g.113062547
059IED50	GPP	European	c.338C>T/c.338C>T; p.Ser113Leu/p.Ser113Leu	chr2:g.113062547/chr2:g.113062547
GBR0002	GPP	European	c.104A>G/c.338C>T; p.Lys35Arg/p.Ser113Leu	chr2:g.113060926/chr2:g.113062547
GYFAP0016	GPP	European	c.338C>T/c.338C>T; p.Ser113Leu/p.Ser113Leu	chr2:g.113062547/chr2:g.113062547
GYFAP0029	GPP	European	c.142C>T/c.338C>T; p.Arg48Trp/p.Ser113Leu	chr2:g.113062150/chr2:g.113062547
GBR0006	GPP	European	c.338C>T/c.338C>T; p.Ser113Leu/p.Ser113Leu	chr2:g.113062547/chr2:g.113062547

SCAR1690	GPP	Asian Other	c.130G>A/c.338C>T; p.Val44Met/p.Ser113Leu	chr2:g.113062138/chr2:g.113062547
SCAR2074	GPP	East Asian	c.338C>T/c.338C>T; p.Ser113Leu/p.Ser113Leu	chr2:g.113062547/chr2:g.113062547
SCAR2548	GPP	East Asian	c.338C>T/-; p.Ser113Leu/-	chr2:g.113062547/-
1GPP1	GPP	East Asian	c.[115+6T>C;227C>T]/-; [p.Arg10Argfs*1; p.Pro76Leu]/-	chr2:g.[113060943;113062235]/-
15GPP1	GPP	Malay	c.115+6T>C/c.115+6T>C; p.Arg10Argfs*1/p.Arg10Argfs*1	chr2:g.113060943/chr2:g.113060943
27GPP1	GPP	Malay	[c.115+6T>C; c.227C>T]/-; [p.Arg10Argfs*1; p.Pro76Leu]/-	chr2:g.[113060943;113062235]/-
40GPP1	GPP	Malay	c.115+6T>C/-; p.Arg10Argfs*1/-	chr2:g.113060943/-
50GPP1	GPP	Malay	c.115+6T>C/c.115+6T>C; p.Arg10Argfs*1/p.Arg10Argfs*1	chr2:g.113060943/chr2:g.113060943
GBR0011	GPP	East Asian	c.338C>T/c.338C>T; p.Ser113Leu/p.Ser113Leu	chr2:g.113062547/chr2:g.113062547
44GPP1	GPP	European	c.115+6T>C/c.338C>T; p.Arg10Argfs*1/p.Ser113Leu	chr2:g.113060943/chr2:g.113062547
55GPP1	GPP	East Asian	c.115+6T>C/c.[115+6T>C;227C>T]; p.Arg10Argfs*1/p.[Arg10Argfs*1;Pro76Leu]	chr2:g.113060943/ chr2:g.[113060943;113062235]
64GPP1	GPP	East Asian	c.115+6T>C/-; p.Arg10Argfs*1/-	chr2:g.113060943/-

74GPP1	GPP	Malay	[c.115+6T>C; c.227C>T]/-; [p.Arg10Argfs*1; p.Pro76Leu]/-	chr2:g.[113060943;113062235]/-
96GPP1	GPP	Malay	c.115+6T>C/c.115+6T>C; p.Arg10Argfs*1/p.Arg10Argfs*1	chr2:g.113060943/chr2:g.113060943
78GPP1	GPP	East Asian	c.115+6T>C/-; p.Arg10Argfs*1/-	chr2:g.113060943/-
82GPP1	GPP	East Asian	c.115+6T>C/c.115+6T>C; p.Arg10Argfs*1/p.Arg10Argfs*1	chr2:g.113060943/chr2:g.113060943
101GPP1	GPP	East Asian	c.115+6T>C/-; p.Arg10Argfs*1/-	chr2:g.113060943/-
103GPP1	GPP	Malay	c.115+6T>C/c.[115+6T>C;227C>T]; p.Arg10Argfs*1/p.[Arg10Argfs*1;Pro76Leu]	chr2:g.113060943/ chr2:g.[113060943;113062235]
IL01	GPP	Malay	c.115+6T>C/c.[115+6T>C;227C>T]; p.Arg10Argfs*1/p.[Arg10Argfs*1;Pro76Leu]	chr2:g.113060943/ chr2:g.[113060943;113062235]
IL02	GPP	Malay	c.115+6T>C/c.115+6T>C; p.Arg10Argfs*1/p.Arg10Argfs*1	chr2:g.113060943/chr2:g.113060943
IL03	GPP	Malay	c.115+6T>C/c.[115+6T>C;227C>T]; p.Arg10Argfs*1/p.[Arg10Argfs*1;Pro76Leu]	chr2:g.113060943/ chr2:g.[113060943;113062235]
IL04	GPP	Malay	c.115+6T>C/c.115+6T>C; p.Arg10Argfs*1/p.Arg10Argfs*1	chr2:g.113060943/chr2:g.113060943
IL06	GPP	East Asian	c.115+6T>C/-; p.Arg10Argfs*1/-	chr2:g.113060943/-
IL07	GPP	East Asian	c.115+6T>C/c.115+6T>C; p.Arg10Argfs*1/p.Arg10Argfs*1	chr2:g.113060943/chr2:g.113060943

IL08	GPP	European	c.115+6T>C/-; p.Arg10Argfs*1/-	chr2:g.113060943/-
117GPP1	GPP	European	c.115+6T>C/c.115+6T>C; p.Arg10Argfs*1/p.Arg10Argfs*1	chr2:g.113060943/chr2:g.113060943
118GPP1	GPP	European	c.115+6T>C/c.115+6T>C; p.Arg10Argfs*1/p.Arg10Argfs*1	chr2:g.113060943/chr2:g.113060943
121GPP1	GPP	East Asian	c.115+6T>C/-; p.Arg10Argfs*1/-	chr2:g.113060943/-
124GPP1	GPP	East Asian	c.338C>T/-; p.Ser113Leu/-	chr2:g.113062547/-
GYFAP0011	GPP	Malay	c.338C>T/-; p.Ser113Leu/-	chr2:g.113062547/-
GBR0016	GPP	Malay	c.338C>T/-; p.Ser113Leu/-	chr2:g.113062547/-
OVS0014	GPP	European	c.338C>T/-; p.Ser113Leu/-	chr2:g.113062547/-
OVS0015	GPP	European	c.338C>T/-; p.Ser113Leu/-	chr2:g.113062547/-
IL13	GPP	European	c.115+6T>C/c.115+6T>C; p.Arg10Argfs*1/p.Arg10Argfs*1	chr2:g.113060943/chr2:g.113060943
OVS0016	GPP	European	c.115+6T>C/c.115+6T>C; p.Arg10Argfs*1/p.Arg10Argfs*1	chr2:g.113060943/chr2:g.113060943
OVS0018	GPP	European	c.115+5G>A /c.338C>T; p.?/p.Ser113Leu	chr2:g.113060942/chr2:g.113062547

141GPP1	GPP	East Asian	c.304C>T/-; p.Arg102Trp/-	chr2:g.113062513/-
144GPP1	GPP	Malay	c.115+6T>C/c.115+6T>C; p.Arg10Argfs*1/p.Arg10Argfs*1	chr2:g.113060943/chr2:g.113060943
146GPP1	GPP	Malay	c.115+6T>C/-; p.Arg10Argfs*1/-	chr2:g.113060943/-
025HHE49	PPP	European	c.338C>T/-; p.Ser113Leu/-	chr2:g.113062547/-
047SJR50	PPP	European	c.227C>T/-; p.Pro76Leu/-	chr2:g.113062235/-
PUS-01	PPP	European	c.338C>T/-; p.Ser113Leu/-	chr2:g.113062547/-
PUS-23	PPP	European	c.338C>T/-; p.Ser113Leu/-	chr2:g.113062547/-
GBR0033	PPP	European	c.338C>T/c.338C>T; p.Ser113Leu/p.Ser113Leu	chr2:g.113062547/chr2:g.113062547
GBR0048	PPP	European	c.338C>T/-; p.Ser113Leu/-	chr2:g.113062547/-
GBR0050	PPP	European	c.338C>T/-; p.Ser113Leu/-	chr2:g.113062547/-
GBR0089	PPP	European	c.338C>T/-; p.Ser113Leu/-	chr2:g.113062547/-
GBR0091	PPP	European	c.338C>T/-; p.Ser113Leu/-	chr2:g.113062547/-

GBR0092	PPP	European	c.338C>T/-; p.Ser113Leu/-	chr2:g.113062547/-
GBR0121	PPP	European	c.338C>T/c.338C>T; p.Ser113Leu/p.Ser113Leu	chr2:g.113062547/chr2:g.113062547
GBR0130	PPP	European	c.338C>T/c.338C>T; p.Ser113Leu/p.Ser113Leu	chr2:g.113062547/chr2:g.113062547
34GPP1	ACH + GPP	East Asian	c.115+6T>C/c.[115+6T>C;227C>T]; p.Arg10Argfs*1/p.[Arg10Argfs*1;Pro76Leu]	chr2:g.113060943/ chr2:g.[113060943;113062235]
79GPP1	ACH + GPP	East Asian	c.115+6T>C/c.115+6T>C; p.Arg10Argfs*1/p.Arg10Argfs*1	chr2:g.113060943/chr2:g.113060943
113GPP1	ACH + GPP	East Asian	c.115+6T>C/c.115+6T>C; p.Arg10Argfs*1/p.Arg10Argfs*1	chr2:g.113060943/chr2:g.113060943
GB40004	GPP + PPP	European	c.338C>T/-; p.Ser113Leu/-	chr2:g.113062547/-
89GPP1	GPP + PPP	Malay	c.115+6T>C/-; p.Arg10Argfs*1/-	chr2:g.113060943/-

[1]Patient is classified as *IL36RN* positive if they carry at least one variant in the examined locus; [2]Patients highlighted in orange have not previously been reported; [3]As per GRCh38 genome build and *IL36RN* transcript NM_012275; [4]When a patient carries both c.115+6T>C and p.Pro76Leu, they are found on the same allele. While in this study a single Sanger sequencing fragment did not span both loci, this allele has been demonstrated in other publications (44,130) and is self-evident in those patients carrying homozygous c.115+6T>C in addition to heterozygous c.227C>T.

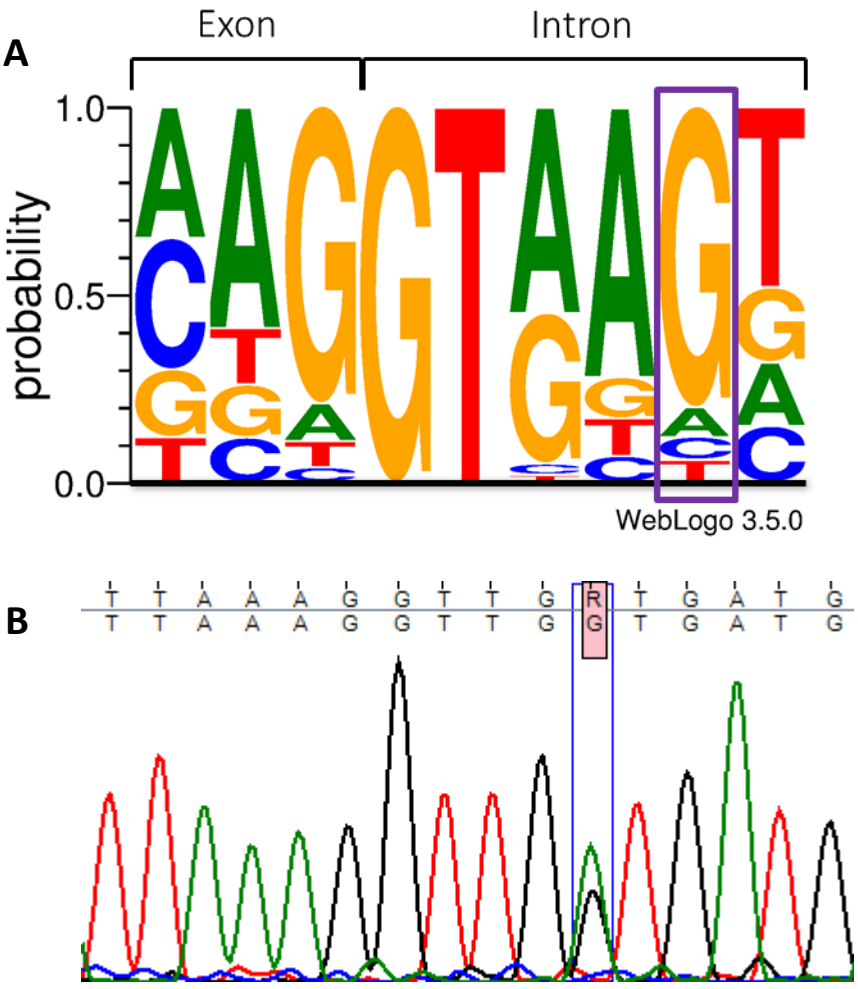


Figure 3.2 Consensus sequence of a donor splice site

A: Sequence logo displaying the frequency of different nucleic acids within a donor splice site. The box highlights the position affected by the c.115+5G>A variant, illustrating that in 80% of splice sites this residue is a guanine. In contrast, adenine is seen in fewer than 10% of sites. Image generated using WebLogo 3 (131,132) with data from (129). B: Chromatogram shows variant as validated by Sanger sequencing.

3.3.2.1 The frequency of *IL36RN* alleles differs between disease subtypes and ethnicities

The prevalence of *IL36RN*-positive patients was much lower in PPP (5.1%) compared to GPP (23.7%) or ACH (17.4%). In addition, GPP and ACH patients were more likely to harbour bi-allelic changes. The prevalence of *IL36RN* variants was therefore significantly lower in PPP (0.03) than GPP (0.19; $P = 1.9x10^{-14}$)) or ACH (0.15; $P = 0.0018$) (Table 3.6). Comparisons involving ACH should however be interpreted with caution, because of the small number of subjects.

As the association between PPP and *IL36RN* variants has previously been questioned (14,41,42), the frequency of the recurrent p.Ser113Leu change was compared in British patients, (the largest geographically homogeneous group within the PPP cohort) vs. population-matched controls. This showed a very significant association between the disease allele and PPP ($P=9.3x10^{-8}$; OR: 10.8; 95% CI: 5.3-22.0) (Table 3.7).

While deleterious *IL36RN* alleles were seen in most ethnic groups, the genetic landscape varied between cohorts of different origin. In the East Asian and Malay datasets, the most common disease allele was the c.115+6T>C splicing variant, with MAFs of 0.238 and 0.098 respectively. In contrast, the most prevalent variant among Europeans was p.Ser113Leu, which had a MAF of 0.29 in our sample. Previous work by the Capon group and others has produced evidence that both variants spread in their respective populations as the result of founder effects (40,44,130,133).

Overall, the highest number of disease alleles was seen in European and East Asian patients (Figure 3.3, panel A). No *IL36RN* variants were detected in South Asian individuals, but this might reflect the small size of this particular cohort (n = 28), as a p.Leu21Pro (c.62T>C) change has been reported in two Pakistani siblings with paediatric GPP (134).

3.3.2.2 *IL36RN* disease alleles are associated with an earlier age of onset

Next, the frequency of *Il36RN* variants were compared in patients with different phenotypic features (Table 3.8), using regression analysis. This method was utilised as the number of ACH cases carrying an *IL36RN* variant was too small for pairwise comparisons with GPP and/or PPP patients to be valid. No significant results were obtained when investigating the relationship between *IL36RN* status and plaque psoriasis concurrence. Conversely, a strong association between *IL36RN* variants and early disease onset was observed (P = 0.003). The presence of a dose-response was also apparent, with patients harbouring biallelic variants tending to develop pustular psoriasis earlier than those with monoallelic variants (Figure 3.3, panel B).

Table 3.6 *IL36RN* variant frequency across disease types

	ACH		GPP		PPP		Multiple diagnoses	
	Positive[1]	Negative	Positive[1]	Negative	Positive[1]	Negative	Positive[1]	Negative
***IL36RN* status (%)**	4 (17.4%)	19 (82.6%)	45 (23.7%)	145 (76.3%)	12 (5.1%)	222 (94.9%)	5 (27.8%)	13 (72.2%)
Prevalence of *IL36RN* alleles (frequency)	7 (0.15)	39 (0.85)	72 (0.19)	308 (0.81)	15 (0.03)	453 (0.97)	8 (0.22)	28 (0.78)

[1]Patients were classified as 'Positive' if they were carrying at least one *IL36RN* variant

Table 3.7 Association between *IL36RN*-p.Ser113Leu and PPP

	p.Ser113Leu	Wild type
Cases[1]	11 (3.6%)	291 (96.4%)
Controls[2]	26 (0.4%)	7402 (99.6%)

[1]British patients only; [2]Control data from UK10K cohort

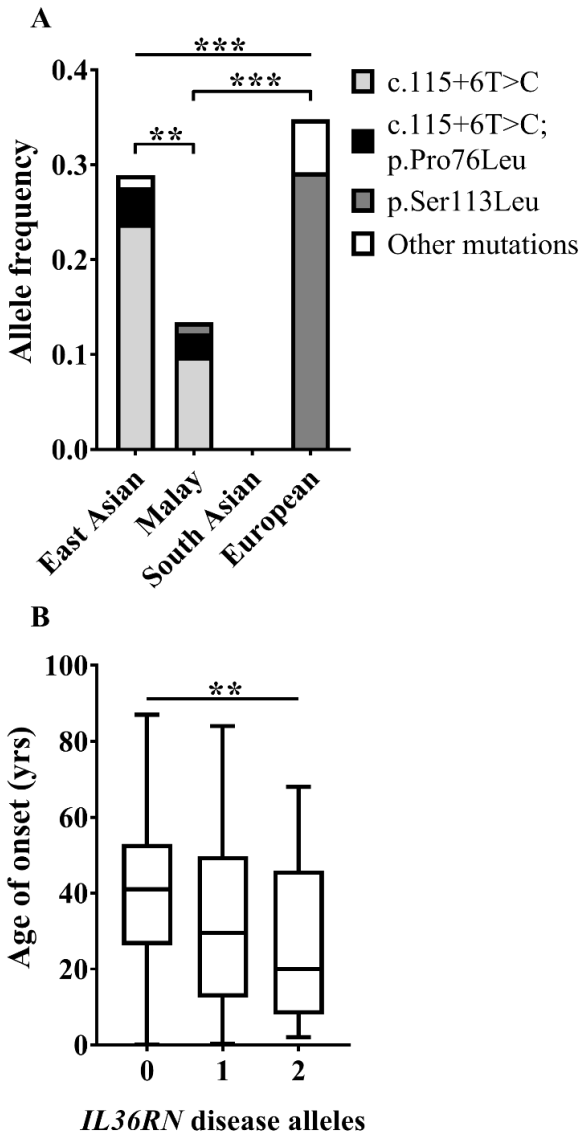


Figure 3.3 Features of *IL36RN* variants within the cohort

(A) There are significant differences in the combined frequency of alleles seen in patients of European, East Asian, Malay and South Asian descent. Comparison of allele frequencies across multiple groups was carried out by chi-squared test. Comparisons between two groups were implemented with Fisher's exact test. Other mutations: alleles seen only once in the cohort. The notation c.115+6T>C;p.Pro76Leu refers to patients carrying the two variants on the same haplotype (44,130). (B) Patients carrying variants in *IL36RN* have earlier disease onset. Statistical significance was assessed by linear regression. **P < 0.01; ***P < 0.001.

Table 3.8 Influence of *IL36RN* genotype on disease phenotype

		Sex		Plaque psoriasis status		Age of onset (yrs)		
Diagnosis	***IL36RN* status**	**Female**	**Male**	**Plaque psoriasis**	**No plaque psoriasis**	**Mean**	**SD**	**Total cases**
ACH	Positive	3 (75.0%)	1 (25.0%)	3 (75.0%)	1 (25.0%)	33.0	27.0	**4**
	Negative	10 (55.6%)	8 (44.4%)	7 (41.2%)	10 (58.8%)	60.3	13.1	**18**
GPP	Positive	30 (66.7%)	15 (33.3%)	18 (42.9%)	24 (57.1%)	27.3	22.0	**45**
	Negative	98 (69.5%)	43 (30.5%)	106 (76.8%)	32 (23.2%)	32.4	18.5	**145**
PPP	Positive	7 (58.3%)	5 (41.7%)	4 (33.3%)	8 (66.7%)	36.6	20.2	**12**
	Negative	163 (75.8%)	52 (24.2%)	51 (26.7%)	140 (73.3%)	43.5	15.0	**217**

3.3.3 Analysis of *AP1S3* disease alleles

To date, only two *AP1S3* disease alleles (p.Phe4Cys (c.11T>G) and p.Arg33Trp (c.97C>T)) have been identified, both of which fall within the second exon of the gene (46,47). Therefore, while 249 patients were screened for the entire coding region, the remaining 74 were only sequenced for exon 2. This identified 24 individuals who carried an *AP1S3* variant (2 ACH, 4 GPP, 14 PPP and 4 with multiple diagnoses) (Table 3.9). In keeping with published findings (41,44,47), these changes were only seen in the heterozygous state and in individuals of European descent. The frequency of disease alleles was similar in GPP (0.05), ACH (0.05) and PPP (0.03) (Table 3.10), although the small number of ACH and GPP cases means caution is required when comparing these groups. Of note, three patients (2 GPP, 1 PPP) carried both *AP1S3* and *IL36RN* disease alleles (Table 3.11).

While the analysis of genotype-phenotype correlations did not yield significant results, it is noteworthy that almost all patients with *AP1S3* variants were female (23/24) (Table 3.12). The distortion of sex ratios does not reach statistical significance (P=0.06), but this may simply reflect the small number of *AP1S3*-positive individuals, raising the possibility that sex-specific factors (such as hormones or X-linked genes) might modify the penetrance of *AP1S3* disease alleles.

Table 3.9 Variants observed in AP1S3 positive patients

Sample ID[1]	Diagnosis	Chromosomal location[2]	*AP1S3* genotype[3] cDNA; protein
OVS0021	ACH	chr2:g.223777776	c.97C>T/-; p.Arg33Trp/-
OVS0024	ACH	chr2:g.223777776	c.97C>T/-; p.Arg33Trp/-
GBR0111	GPP	chr2:g.223777776	c.97C>T/-; p.Arg33Trp/-
SCAR1690	GPP	chr2:g.223777862	c.11T>G/-; p.Phe4Cys/-
OVS0014	GPP	chr2:g.223777862	c.11T>G/-; p.Phe4Cys/-
GYFAP0163	GPP	chr2:g.223777776	c.97C>T/-; p.Arg33Trp/-
025HHE49	PPP	chr2:g.223777862	c.11T>G/-; p.Phe4Cys/-
100IRU41	PPP	chr2:g.223777862	c.11T>G/-; p.Phe4Cys/-
PUS-06	PPP	chr2:g.223777776	c.97C>T/-; p.Arg33Trp/-
GBR0029	PPP	chr2:g.223777776	c.97C>T/-; p.Arg33Trp/-
GBR0052	PPP	chr2:g.223777776	c.97C>T/-; p.Arg33Trp/-
GBR0056	PPP	chr2:g.223777862	c.11T>G/-; p.Phe4Cys/-
GBR0073	PPP	chr2:g.223777776	c.97C>T/-; p.Arg33Trp/-
GBR0078	PPP	chr2:g.223777862	c.11T>G/-; p.Phe4Cys/-
GBR0082	PPP	chr2:g.223777862	c.11T>G/-; p.Phe4Cys/-
GBR0122	PPP	chr2:g.223777776	c.97C>T/-; p.Arg33Trp/-
GYFAP0019	PPP	chr2:g.223777862	c.11T>G/-; p.Phe4Cys/-
GBR0140	PPP	chr2:g.223777776	c.97C>T/-; p.Arg33Trp/-
GBR0149	PPP	chr2:g.223777776	c.97C>T/-; p.Arg33Trp/-
P020022	PPP	chr2:g.223777776	c.97C>T/-; p.Arg33Trp/-
GYFAP0014	ACH + GPP	chr2:g.223777776	c.97C>T/-; p.Arg33Trp/-
GBR0007	ACH + GPP	chr2:g.223777776	c.97C>T/-; p.Arg33Trp/-
GBR0126	GPP + PPP	chr2:g.223777862	c.11T>G/-; p.Phe4Cys/-
GBR0009	GPP + PPP	chr2:g.223777862	c.11T>G/-; p.Phe4Cys/-

[1]Patients highlighted in orange have not previously been reported; [2]As per GRCh38 genome build; [3]*AP1S3* transcript NM_001039569

Table 3.10 *AP1S3* variant frequency across disease types[1]

	ACH		GPP		PPP		Multiple diagnoses	
	Positive[2]	Negative	Positive[2]	Negative	Positive[2]	Negative	Positive[2]	Negative
***AP1S3* status (%)**	2 (10.5%)	17 (89.5%)	4 (10.8%)	33 (89.2%)	14 (6.6%)	198 (93.4%)	4 (36.4 %)	7 (63.6%)
Prevalence of *AP1S3* alleles (frequency)	2 (0.05)	36 (0.95)	4 (0.05)	70 (0.95)	14 (0.03)	410 (0.97)	4 (0.18)	18 (0.82)

[1]The p.Phe4Cys and p.Arg33Trp variants have no frequency in East-Asian populations and therefore were not screened in patients from this ethnic group. [2]Patients were classified as 'Positive' if they were carrying at least one variant at the examined locus

Table 3.11 Digenic variants observed within the cohort

Sample ID	Diagnosis	***IL36RN* genotype** **cDNA; protein**	***AP1S3* genotype** **cDNA; protein**
SCAR1690	GPP	c.338C>T/c.130G>A; p.Val44Met/p.Ser113Leu	c.11T>G/-; p.Phe4Cys/-
OVS0014	GPP	c.338C>T/-; p.Ser113Leu/-	c.11T>G/-; p.Phe4Cys/-
ERS025[1]	PPP	c.338C>T/-; p.Ser113Leu/-	c.11T>G/-; p.Phe4Cys/-

[1]Patients highlighted in orange have not previously been reported

Table 3.12 Influence of *AP1S3* genotype on disease phenotype

		Sex		Plaque psoriasis status		Age of onset (yrs)		
Diagnosis	***AP1S3* status**	**Female**	**Male**	**Plaque psoriasis**	**No plaque psoriasis**	**Mean**	**SD**	**Total cases**
ACH	Positive	2 (100%)	0 (0%)	2 (100%)	0 (0.0%)	65.0	15.6	**2**
	Negative	11 (64.7%)	6 (35.3%)	6 (40.0%)	9 (60.0%)	55.6	16.4	**17**
GPP	Positive	4 (100%)	0 (0%)	1 (50.0%)	1 (50.0%)	39.7	37.8	**4**
	Negative	23 (71.9%)	9 (28.1%)	10 (34.5%)	19 (65.5%)	30.2	26.6	**33**
PPP	Positive	13 (92.9%)	1 (7.1%)	4 (30.8 %)	9 (69.2%)	48.1	11.6	**14**
	Negative	149 (75.3%)	49 (24.7%)	43 (23.8%)	138 (76.2%)	43.8	15.1	**198**

3.3.4 *CARD14* and pustular psoriasis

Genetic studies of *CARD14* have identified a clear disease allele hotspot spanning exons 3 and 4 of the gene (40,48,51). The full coding region (21 exons) was therefore only screened in 106 patients, whereas the remaining were only sequenced for exons 3 and 4.

In addition to three previously reported Chinese cases carrying a heterozygous p.Asp176His (c.526G>C) change (48), five British PPP patients were found to harbour rare variants in the gene (Table 3.13, Figure 3.4). These alleles were predicted to have damaging potential by online tools. Interestingly, three of the four variants were not located within the hotspot, illustrating the importance of continuing to explore regions outside of those of established interest. While six of the eight patients also suffered from plaque psoriasis, the small size of the dataset meant there was insufficient power to detect genotype-phenotype correlations.

Table 3.13 Details of *CARD14* alleles with deleterious potential

Sample ID[1]	Diagnosis	Ethnicity	Chromosomal location[2]	*CARD14* genotype[3] cDNA; protein	CADD score[4]
16GPP1	GPP	Chinese	chr17:g.80184089	c.526G>C/-; p.Asp176His/-	26.7
81GPP1	GPP	Chinese	chr17:g.80184089	c.526G>C/-; p.Asp176His/-	
108GPP1	GPP	Chinese	chr17:g.80184089	c.526G>C/-; p.Asp176His/-	
GYFAP0018	PPP	European	chr17:g.80198512	c.1772C>T/-; p.Thr591Met/-	28.1
GYFAP0023	PPP	Unknown	chr17:g.80184107	c.544C>T/-; p.Arg182Cys/-	24.3
POPLM0002	PPP	European	chr17:g.80182675	c.234G>T/-; p.Lys78Asn/-	28.2
POPLM0008	PPP	European	chr17:g.80202245	c.2044C>T/-; p.Arg682Trp/-	35
PLPLM0003	PPP	European	chr17:g.80202245	c.2044C>T/-; p.Arg682Trp/-	

[1]Patients highlighted in orange have not previously been reported; [2]As per GRCh38 genome build; [3]*CARD14* transcript NM_024110;

[4]A CADD score greater than 15 is indicative of pathogenicity

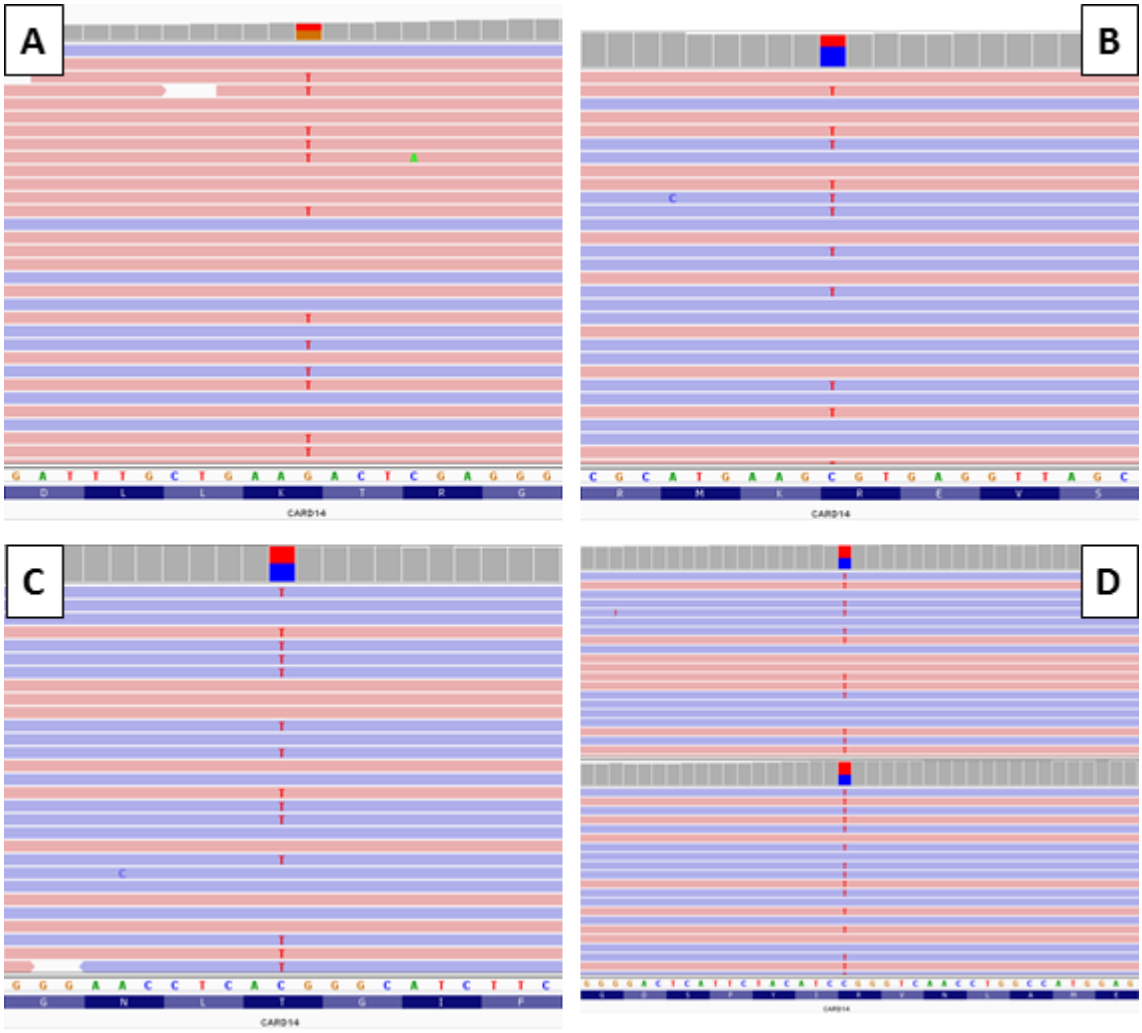


Figure 3.4 IGV visualisations of *CARD14* variants

Screenshots of IGV visualisation of variants seen in *CARD14* by whole exome sequencing.

A: c.234G>T; p.Lys78Asn. **B**: c.544C>T; p.Arg182Cys. **C**: c.1772C>T; p.Thr591Met. **D**: c.2044C>T; p.Arg682Trp. The recurrent c.526G>C; p.Asp176His variant is not shown as this is well established as disease associated.

3.4 Discussion

This work aimed to analyse a large, multi-ethnic pustular psoriasis cohort for the purposes of identifying demographic, clinical and genetic distinctions between the disease subtypes. Previous studies have generally examined patients of only one ethnicity and/or been restricted by smaller cohort sizes (28,41,42,44,46,48).

A high degree of heterogeneity was present within our dataset. However, there were clear differences between features of GPP and PPP. Generalised pustular psoriasis patients are on average 13 years younger when the disease manifests than those with PPP, and paediatric onset is not uncommon. Although both disease subtypes displayed a sex ratio skewed towards females, the bias was much greater in PPP. In contrast, while higher than the general population, the prevalence of plaque psoriasis was lower in PPP than GPP. Finally, an extremely high proportion of PPP patients had a history of smoking, and this was associated with increased rates of concurrent plaque psoriasis.

The analysis of patient clinical and demographic characteristics therefore indicates that PPP largely affects middle-aged women with a history of smoking. These features have been reported by others (18,21,22,126,135), but this is the first demonstration that they are specific to PPP and differentiate the disease from other forms of pustular psoriasis.

It is widely known that there are differences in the immune response that correlate with sex: females mount a stronger challenge to pathogens and also account for 80% of autoimmune disease cases (136,137). One of the more extreme sex biases is seen in systemic lupus erythematosus (SLE), where 90% of sufferers are female. A recent study has revealed that the gene encoding toll-like receptor 7 (*TLR7*), which drives a key pathway in SLE pathogenesis, escapes X chromosome inactivation in immune cells (138) and is therefore more active in

females. Meanwhile, transcriptomic analyses of human skin have identified the transcription factor *VGLL3* (which is itself expressed more highly in females) as a key regulator of other female-biased targets that are associated with autoimmune diseases (139). Finally, examples of hormonal effects on the immune system include the numerous aspects of immunity that are affected by oestradiol, including monocyte differentiation, neutrophil proliferation and expression of pro-inflammatory cytokines (137). It is therefore unsurprising to find a sex bias in pustular psoriasis, but in-depth functional studies will be needed to determine whether this phenomenon is mediated by any of the general mechanisms described above (skewed inactivation of X-linked genes, sex-specific differences in gene expression, hormonal influences).

One hypothesis to explain the relationship between smoking and PPP is that the disease is the result of an excessive response to nicotine (127,140). To explore this possibility, a study examined the expression and distribution of nicotinic acetylcholine receptors (nAChRs, of which nicotine is an agonist) in the skin of healthy non-smokers, healthy smokers and PPP patients. Both α3 and α7 nAChRs exhibited abnormal distribution and elevated expression in lesional PPP skin. In addition, healthy smokers had higher nAChR levels than healthy non-smokers and atypical patterns of α7 nAChR staining were seen in non-lesional patient biopsies. This led the authors to suggest that smoking modulated nAChR expression and that the potential for abnormal nicotine responses exists in the skin of patients prior to the onset of inflammation (140).

More recently, a link between smoking, the aryl hydrocarbon receptor (AhR) and psoriatic inflammation has emerged. Signalling via AhR downregulates expression of genes related to psoriasis, reducing levels of inflammation (141). The region encoding the repressor of the AhR pathway (AHRR) contains multiple CpG islands that are differentially methylated in smokers. In particular, smoking causes AHRR to become hypomethylated, which results in an increase in

gene expression. This is likely to downregulate the AhR pathway and impair its anti-inflammatory function in skin (142,143).

It is notable that PPP also appears distinct from GPP and ACH at the genetic level, with a much lower prevalence of *IL36RN* variants. Evidence for divergent immunopathogeneses was also observed in a study that profiled the transcriptome of GPP and PPP lesions, vs. that of normal skin. This revealed a number of differentially expressed genes (DEG) that were unique to each disease. Of note GPP-specific DEGs tended to cluster to very specific cellular processes (e.g. protein ubiquitination), while PPP-only DEGs showed significant enrichment for pathways related to T-cell signalling (144). Thus, PPP may be a disease driven by a dysfunction of adaptive immunity, caused by abnormalities in cytokines other than IL-36. Given that the IL-36 pathway has also been implicated in the pathogenesis of plaque psoriasis (25,33), it is possible that the lower rates of this condition in PPP may be linked to the relative rarity of *IL36RN* disease alleles.

The association between *IL36RN* alleles and earlier disease onset has been previously reported in GPP (15) but here is robustly demonstrated in all disease forms. This work therefore suggests that patients with symptoms that appear before the age of 30 (GPP) or 40 (PPP/ACH) should be prioritised for *IL36RN* screening, in addition to GPP patients presenting with systemic inflammation (15).

While it was not possible to detect statistically robust correlations between *AP1S3* genotype and patient phenotype, an interesting trend was identified: almost all carriers of *AP1S3* disease alleles are female. The screening of additional patients may help to bring statistical significance to this observation, which would then merit investigation into the sex-specific modifiers underlying the association. As described above, these could take the form of hormonal effects or interactions with proteins with sex-biased expression.

It was also difficult to draw conclusions from the results of the *CARD14* screening, but a number of new alleles with pathogenic potential were identified. One of these variants (c.544C>T; p.Arg182Cys) lies within a known disease allele hotspot, in a region that encodes a coiled-coil domain. This has an autoinhibitory function, as it maintains the protein in a close conformation that prevents its oligomerisation in the absence of inflammatory stimuli. Disease alleles disrupt the regulatory function of the coiled-coil domain, leading to constitutive protein oligomerisation, which in turn causes abnormal NF-kB activation (48). However, other changes observed in this study were located outside of the coiled-coil region and so further work would be needed to experimentally validate their pathogenic potential. One variant, p.Lys78Asn (c.234G>T), maps to the caspase recruitment domain (CARD) region, while the remaining two (c.1772C>T; p.Thr591Met and c.2044C>T; p.Arg682Trp) are within a PDZ domain (145,146). Interestingly, a study found that, in *CARD11*, deletion of the wider region containing the PDZ domain resulted in elevated NFκB signalling (147).

More broadly, the genetic analysis of the cohort emphasises the fact that known disease genes account for only a fifth of pustular psoriasis cases. There is therefore still a real need to identify new genes involved in disease pathogenesis, particularly in PPP where fewer than 10% of individuals harbour a known disease allele.

Key to expediting the process of gene identification is the standardised ascertainment of patients, a task that has been facilitated by the work of the ERASPEN consortium (1,148). This work expands on the consensus diagnostic criteria outlined by ERASPEN and further demonstrates the value of collaborative approaches for the rigorous study of rare diseases.

4 Improving the classification of *IL36RN* variants

4.1 The variant landscape of *IL36RN*

In the seven years since *IL36RN* variants were first associated with pustular psoriasis, patient screening has revealed additional variants in the gene. As of October 2018, there are 20 alleles listed on Infevers (a registry of variants associated with hereditary auto-inflammatory disorders) (39), with an additional five awaiting upload. There are in total 18 missense substitutions, four stop-gain variants, one splicing change, and two deletions (Table 4.1). However, as the only requirement for inclusion on Infevers is that the variant has been seen in a patient, it is possible that there are benign variants included in the registry. Indeed, one allele listed on Infevers (c.140A>G; p.Asn47Ser) has a frequency of 5.4% in East Asians, suggesting it is unlikely to be disease-causing. This highlights the importance of examining population-specific data.

It is not always practical for functional work to be promptly carried out to assess the effect of a new nucleotide change. Moreover, there is not a single *in-vitro* assay that has been applied to the analysis of all known variants and is considered as a gold standard in the field. In these circumstances, researchers and clinicians generally turn to one or more of a diverse range of *in-silico* pathogenicity prediction tools. However, different algorithms have been used to assess the effects of different variants.

This lack of consistency has important clinical implications, as *IL36RN* screening is increasingly used as a diagnostic test for generalised pustular psoriasis. Furthermore, it is likely that as the gene is sequenced in more patients, additional variants will be uncovered. There is therefore a clear need to achieve a better understanding of *IL36RN* alleles and their effects on protein function. Here, this goal was pursued by attempting a systematic characterisation of *IL36RN* missense variants, based on the integration of *in-vitro* and *in-silico* approaches.

Table 4.1 Published variants in *IL36RN*

Variant[1]	Type	ExAC MAF (min – max)[2]	Original publication	Chromosomal location; cDNA change[3]
p.Val2Phe	Missense	NA[4]	Bal et al (149)	chr2:g.113059442; c.4G>T
p.Arg10X	Stop-gain	0 – 1.156×10^{-4}	Sugiura et al (150)	chr2:g.113059466; c.28C>T
p.Ser14X	Stop-gain	NA	Tauber et al (151)	chr2:g.113060863; c.41C>A
p.Leu21Pro	Missense	NA	Ellingford et al (134)	chr2:g.113060884; c.62T>C
p.Leu27Pro	Missense	0 – 0.002611	Marrakchi et al (29)	chr2:g.113060902; c.80T>C
p.His32Arg	Missense	NA	Korber et al (152)	chr2:g.113060917; c.95A>G
p.Lys35Arg	Missense	0 – 4.97×10^{-4}	Setta-Kaffetzi et al (44)	chr2:g.113060926; c.104A>G
c.115+5G>A	Splicing motif disrupted	0 – 6.046×10^{-5}	This study (section 3.3)	chr2:g.113060942; c.115+5G>A
c.115+6T>C	Splicing motif disrupted	0 – 0.01383	Farooq et al (133)	chr2:g.113060943; c.115+6T>C
p.Ile42Asn	Missense	0 – 1.528×10^{-5}	Takeichi et al (153)	chr2:g.113062133; c.125T>A
p.Val44Met	Missense	NA	Mössner et al (41)	chr2:g.113062138; c.130G>A
p.Asn47Ser	Missense	8.731×10^{-5} – 0.05405	Li et al (130)	chr2:g.113062148; c.140A>G
p.Arg48Trp	Missense	0 – 0.001143	Onoufriadis et al (28)	chr2:g.113062150; c.142C>T
p.Pro76Leu	Missense	0 – 0.001157	Korber et al (152)	chr2:g.113062235; c.227C>T

p.Pro82Leu	Missense	0 – 0.004636	Takahashi et al (43)	chr2:g.113062454; c.245C>T
p.Glu94X	Stop-gain	NA	Korber et al (152)	chr2:g.113062489; c.280G>T
c.295_300del	Non-frameshift deletion	NA	Mössner et al (41)	chr2:g.113062504_113062508; c.295_300del
p.Arg102Trp	Missense	0 – $2.313x10^{-4}$	Setta-Kaffetzi et al (44)	chr2:g.113062513; c.304C>T
p.Arg102Gln	Missense	0 – $1.501x10^{-5}$	Li et al (130)	chr2:g.113062514; c.305G>A
p.Arg103Gln	Missense	0 – $1.515x10^{-4}$	Mössner et al (41)	chr2:g.113062517; c.308G>A
p.Glu112Lys	Missense	0 – $1.502x10^{-5}$	Hayashi et al (154)	chr2:g.113062543; c.334G>A
p.Ser113X	Stop-gain	0 – $1.502x10^{-5}$	Mössner et al (41)	chr2:g.113062547; c.338C>A
p.Ser113Leu	Missense	0 – 0.006505	Onoufriadis et al (28)	chr2:g.113062547; c.338C>T
p.Thr123Arg	Missense	NA	Farooq et al (133)	chr2:g.113062577; c.368C>G
p.Thr123Met	Missense	0 – $6.056x10^{-5}$	Kanazawa et al (155)	chr2:g.113062577; c.368C>T
c.420_426del	Frameshift deletion	NA	Arostegui et al (156)	chr2:g.113062629_ chr2:g.113062634; c.420_426del

[1]Name most commonly used in literature; [2]Smallest and largest minor allele frequencies, taken from the Exome Aggregation Consortium; [3]As per GRCh38 genome build and *IL36RN* transcript NM_012275; [4]NA: not available (variant has not been seen in any ExAC individual)

4.2 Surveying the effects of *IL36RN* variants

While the splicing, deletion and stop-gain alleles found on Infevers are very likely to have deleterious effects on the IL-36Ra protein, missense changes are not so easily classified. This study therefore assessed all 18 missense variants that have been observed in pustular psoriasis cases (patient set, Table 4.2). To further explore the relationship between *in-silico* predictions and functional impact, a further 12 non-synonymous SNVs were selected, so as to represent variation that affects various protein segments and is associated with a broad range of CADD scores (population set, Table 4.3).

Table 4.2 *IL36RN* variants selected for study: patient set

Variant	rs ID	Location; cDNA change[1]
p.Val2Phe	rs1292126146	chr2:g.113059442; c.4G>T
p.Leu21Pro	rs1285007154	chr2:g.113060884; c.62T>C
p.Leu27Pro	rs387906914	chr2:g.113060902; c.80T>C
p.His32Arg	NA	chr2:g.113060917; c.95A>G
p.Lys35Arg	rs187015338	chr2:g.113060926; c.104A>G
p.Ile42Asn	rs775349262	chr2:g.113062133; c.125T>A
p.Val44Met	rs776622427	chr2:g.113062138; c.130G>A
p.Asn47Ser	rs28938777	chr2:g.113062148; c.140A>G
p.Arg48Trp	rs151325121	chr2:g.113062150; c.142C>T
p.Pro76Leu	rs139497891	chr2:g.113062235; c.227C>T
p.Pro82Leu	rs144182857	chr2:g.113062454; c.245C>T
p.Arg102Trp	rs199932303	chr2:g.113062513; c.304C>T
p.Arg102Gln	rs371819085	chr2:g.113062514; c.305G>A
p.Arg103Gln	rs542606182	chr2:g.113062517; c.308G>A
p.Glu112Lys	rs143724424	chr2:g.113062543; c.334G>A
p.Ser113Leu	rs144478519	chr2:g.113062547; c.338C>T
p.Thr123Arg	rs397514629	chr2:g.113062577; c.368C>G
p.Thr123Met		chr2:g.113062577; c.368C>T

[1]As per GRCh38 genome build and *IL36RN* transcript NM_012275

Table 4.3 *IL36RN* variants selected for study: population set

Variant	rs ID	Location; cDNA change[1]	ExAC MAF (min – max)[2]	CADD Score[3]
p.Gln25Arg	rs867378394	chr2:g.113060896; c.74A>G	NA[4]	23.4
p.Glu41Gln	rs771984756	chr2:g.113062129; c.121G>C	0 – $1.1174x10^{-4}$	23.5
p.Ala52Thr	rs755465505	chr2:g.113062162; c.154G>A	0 – $1.158x10^{-4}$	16.7
p.Cys70Arg	rs375718709	chr2:g.113062216; c.208T>C	0 – $1.926x10^{-4}$	24.0
p.Thr77Ile	rs372880215	chr2:g.113062238; c.230C>T	0 - 0.001114	1.2
p.Lys93Glu	rs746109701	chr2:g.113062486; c.277A>G	0 - $1.156x10^{-4}$	0.07
p.Tyr101Phe	rs769214649	chr2:g.113062511; c.302A>T	0 – $3.853x10^{-4}$	12.3
p.Glu138Lys	rs750580815	chr2:g.113062621; c.412G>A	0 - $1.505x10^{-5}$	9.5
p.Gly141Asp	rs758533837	chr2:g.113062631; c.422G>A	0 - $1.508x10^{-5}$	0.03
p.Ala144Thr	rs780261792	chr2:g.113062639; c.430G>A	0 - $6.047x10^{-5}$	0.17
p.Ile146Val	rs202059991	chr2:g.113062645; c.436A>G	0 – $2.576x10^{-4}$	13.0
p.Gln153Arg	rs771496493	chr2:g.113062667; c.458A>G	0 - $1.559x10^{-5}$	9.5

[1]As per GRCh38 genome build and *IL36RN* transcript NM_012275; [2]Smallest and largest minor allele frequencies, taken from the Exome Aggregation Consortium; [3]Correct as of Jan 2018; [4]NA: not available (variant has not been seen in any ExAC individual)

4.2.1 *In-silico* pathogenicity predictions

Table 4.4 and Table 4.5 display the output for all assessed variants from three commonly used pathogenicity prediction tools (correct as of October 2018). These were selected to ensure there was no overlap in the annotations used in the respective training sets. Two popular *in-silico* predictors (SIFT and PolyPhen-2) are not shown here, because they formed part of the training data for CADD (96). It should be noted that the work in the following chapter was carried out prior to the analysis described here, and utilised aspects of a pre-established pipeline where annotations from SIFT, PolyPhen-2 and CADD were examined in parallel. As such, the variant scoring system outlined in Materials and Methods (Table 2.12) does not apply here.

The most conservative tool was PROVEAN, which classified most variants (8/18 in the patient set and 11/12 in the population set) as 'Neutral'. In contrast, CADD predicted that virtually all changes (18/18 in the patient set and 9/12 in the population set) would have an effect on protein function. The output of MutationTaster was between that of PROVEAN and CADD with half of the changes labelled as "Polymorphism" (5/18 in the patient set and 10/12 in the population set) and the other half as "Disease Causing". As a result of these discrepancies, the majority of variants (8/18 in the patient set and 8/12 in the population set) were associated with discordant predictions, highlighting the need for functional characterisation.

Table 4.4 Pathogenicity predictions for variants in the patient set

Variant	PROVEAN	MutationTaster	CADD[1]
c.4G>T; p.Val2Phe	Neutral	Disease causing	25.4
c.62T>C; p.Leu21Pro	Deleterious	Disease causing	27.3
c.80T>C; p.Leu27Pro	Deleterious	Disease causing	23.3
c.95A>G; p.His32Arg	Neutral	Polymorphism	19.17
c.104A>G; p.Lys35Arg	Neutral	Disease causing	22.8
c.125T>A; p.Ile42Asn	Deleterious	Disease causing	26.9
c.130G>A; p.Val44Met	Neutral	Polymorphism	18.35
c.140A>G; p.Asn47Ser	Neutral	Polymorphism	22.2
c.142C>T; p.Arg48Trp	Deleterious	Disease causing	25.4
c.227C>T; p.Pro76Leu	Deleterious	Disease causing	22.7
c.245C>T; p.Pro82Leu	Neutral	Disease causing	21.6
c.304C>T; p.Arg102Trp	Deleterious	Disease causing	22.9
c.305G>A; p.Arg102Gln	Neutral	Polymorphism	25.2
c.308G>A; p.Arg103Gln	Neutral	Polymorphism	21.5
c.334G>A; p.Glu112Lys	Deleterious	Disease causing	27.1
c.338C>T; p.Ser113Leu	Deleterious	Disease causing	22.8
c.368C>G; p.Thr123Arg	Deleterious	Disease causing	23.4
c.368C>T; p.Thr123Met	Deleterious	Disease causing	23.6

[1]CADD scores>15 are considered as evidence of pathogenicity

Table 4.5 Pathogenicity predictions for variants in population set

Variant	PROVEAN	MutationTaster	CADD[1]
c.74A>G; p.Gln25Arg	Neutral	Polymorphism	23
c.121G>C; p.Glu41Gln	Neutral	Disease causing	24.5
c.154G>A; p.Ala52Thr	Neutral	Polymorphism	19.1
c.208T>C; p.Cys70Arg	Deleterious	Disease causing	23.9
c.230C>T; p.Thr77Ile	Neutral	Polymorphism	10.63
c.277A>G; p.Lys93Glu	Neutral	Polymorphism	7.538
c.302A>T; p.Tyr101Phe	Neutral	Polymorphism	18.54
c.412G>A; p.Glu138Lys	Neutral	Polymorphism	11.99
c.422G>A; p.Gly141Asp	Neutral	Polymorphism	1.371
c.430G>A; p.Ala144Thr	Neutral	Polymorphism	3.766
c.436A>G; p.Ile146Val	Neutral	Polymorphism	18.7
c.458A>G; p.Gln153Arg	Neutral	Polymorphism	18.36

[1]CADD scores>15 are considered as evidence of pathogenicity; here CADD scores between 10 and 15 were classed as intermediate in order to avoid missing variants of moderate effect.

4.2.2 Protein expression

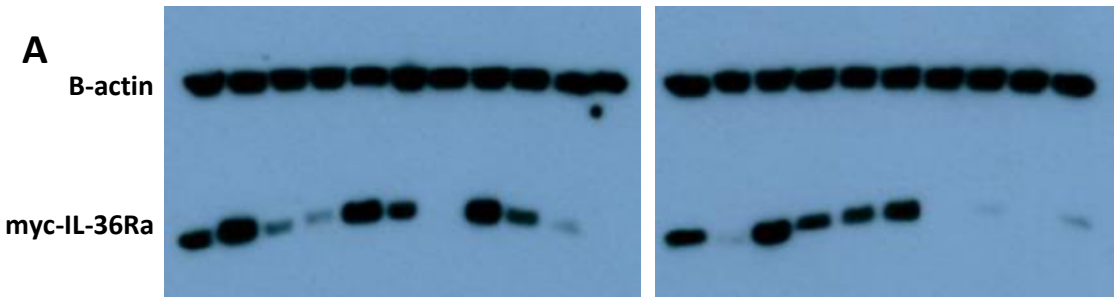
As missense changes may affect protein expression, the results of pathogenicity predictions were followed up by investigating the effects of *IL36RN* variants on IL-36Ra expression. HeLa cells were transfected with wild-type and myc-tagged *IL36RN* constructs and the lysates were analysed by western blotting, using an anti-myc antibody. In this system, the intensity of the band visible on the Western blot serves as a proxy for protein expression.

The assay confirmed that well-characterised disease-causing alleles (p.Leu27Pro and p.Ser113Leu) result in a marked reduction in protein expression. Conversely, the p.Asn47Ser variant, which has high frequency in East Asian populations, did not appear to affect IL-36Ra levels. Thus, control samples behaved as expected.

A third (5/15) of the remaining missense changes from the patient set exhibited reduced expression (p.Ile42Asn, p.Pro76Leu, p.Glu112Lys, p.Thr123Arg, p.Thr123Met; Figure 4.1).

Many of the alleles from the population set had little effect on protein expression (Figure 4.2). There were, however, some exceptions. The p.Glu41Gln and p.Glu138Lys, p.Gly141Asp changes were associated with reduced IL-36Ra levels and p.Cys70Arg severely disrupted expression, such that no protein was visible in the blot.

Thus, reduced protein expression is a relatively common effect of *IL36RN* alleles. However, there are other mechanisms whereby amino acid changes can affect protein activity. For example, it has been shown that the p.Val2Phe change disrupts the processing step which activates IL-36Ra (149). Similarly, it has been suggested that p.His32Arg may impair the interaction between IL-36Ra and the IL-36 receptor (151). It was therefore important to examine whether the variant proteins could function normally *in-vitro*.



L-R: Wild type, p.Val2Phe, p.Leu21Pro, p.Leu27Pro, p.His32Arg, p.Lys35Arg, p.Ile42Asn, p.Val44Met, p.Asn47Ser, p.Arg48Trp, untransfected control

L-R: Wild type, p.Pro76Leu, p.Pro82Leu, p.Arg102Trp, p.Arg102Gln, p.Arg103Gln, p.Glu112Lys, p.Ser113Leu, p.Thr123Arg, p.Thr123Met

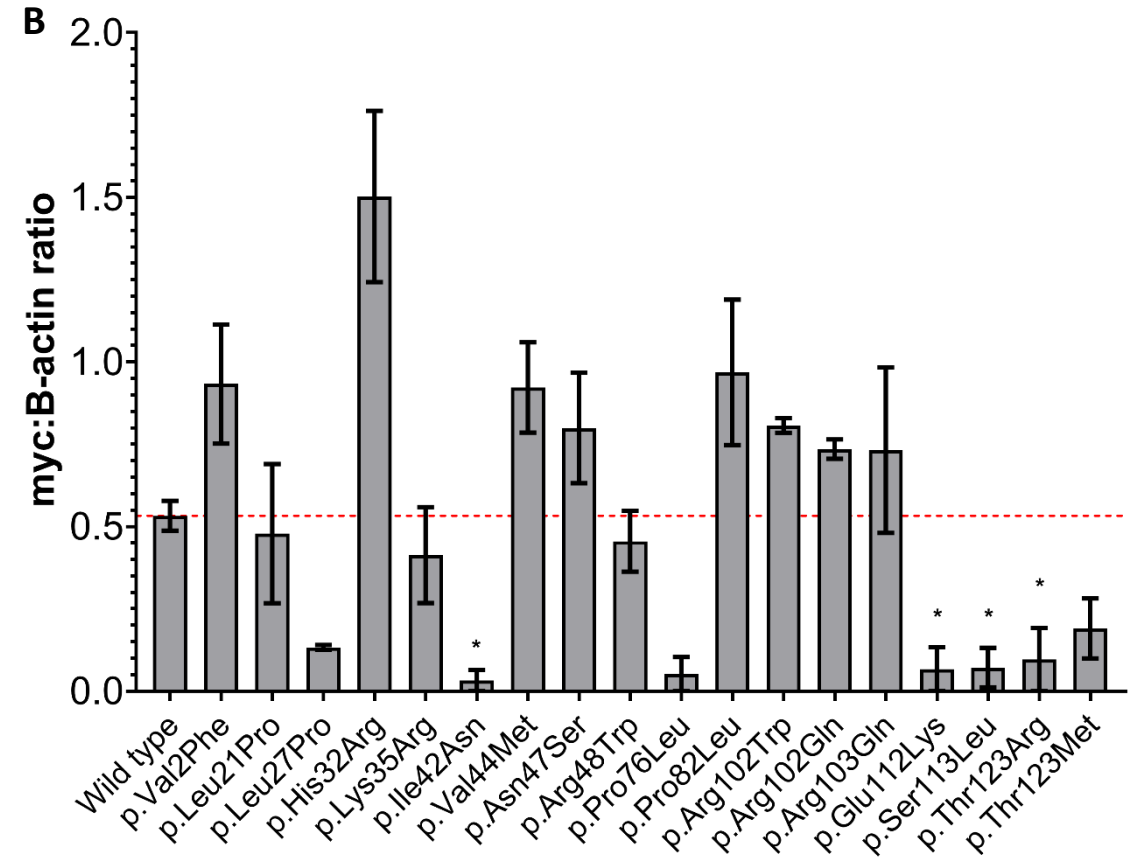


Figure 4.1 The expression of variant forms of IL-36Ra, patient set

(A) Image displays representative Western blot, with variants in order seen in graph. (B) Densitometry measurements were derived with ImageJ and the ratio of myc-tagged IL-36Ra to β-actin was calculated for each lane. Plots show mean ratios observed in at least two independent transfections. Error bars show standard error; the red line indicates the mean ratio for wild type IL-36Ra. The mean ratio for each variant was compared to wild type by Dunnett's multiple comparisons test. * $P < 0.05$

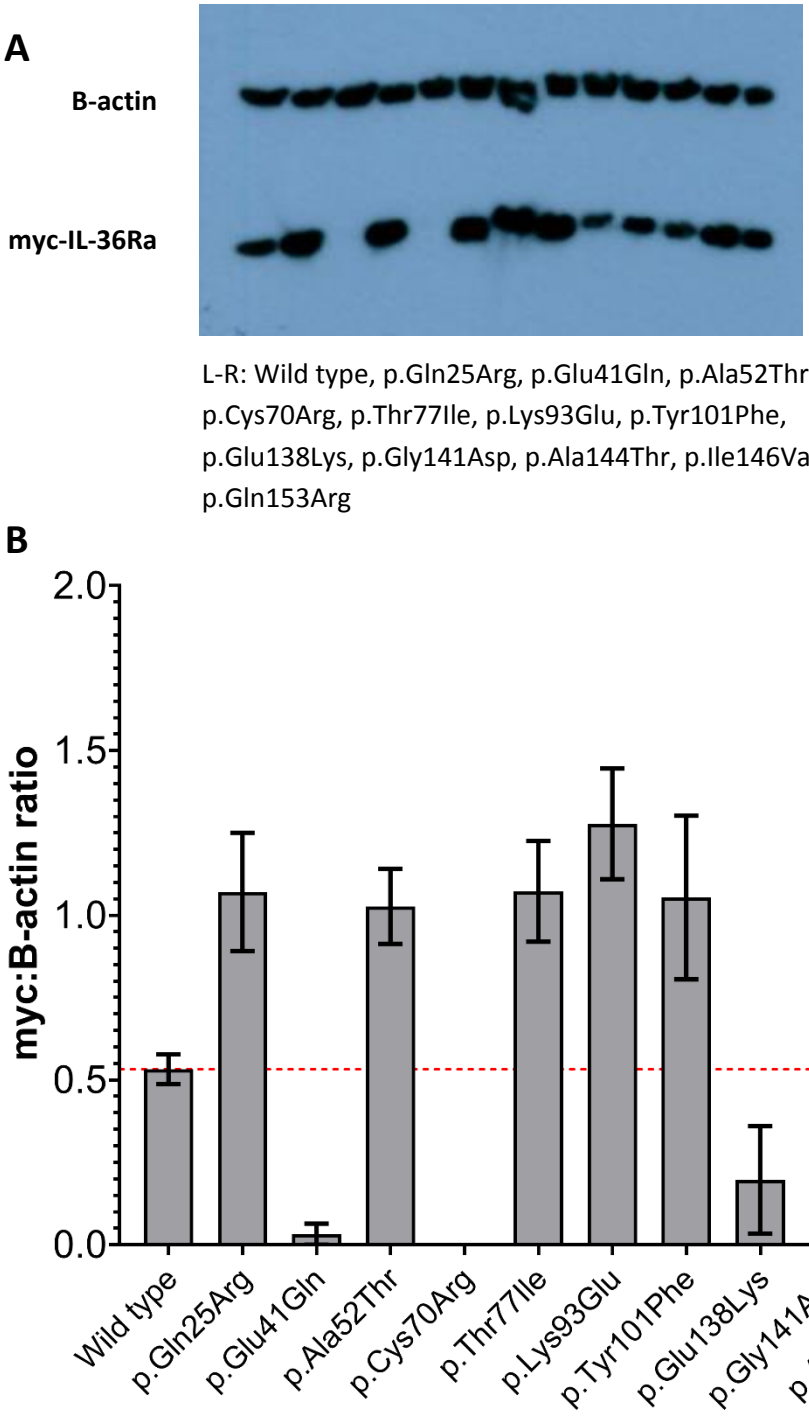


Figure 4.2 The expression of variant forms of IL-36Ra, population set

(A) Image displays representative Western blot, with variants in order seen in graph. (B) Densitometry measurements were derived with ImageJ and the ratio of myc-tagged IL-36Ra to β-actin was calculated for each lane. Plots show mean ratios observed in at least two independent transfections. Error bars show standard error; the red line indicates the mean ratio for wild type IL-36Ra. The mean ratio for each variant was compared to wild type by Dunnett's multiple comparisons test.

4.2.3 Protein activity

The biological activity of a subset of IL-36Ra alleles was investigated by assaying the effects of wild type or variant IL-36Ra on IL-36 signal transduction. The examined variants included three that were predicted to be pathogenic and were found to reduce protein expression (p.Leu27Pro, p.Pro76Leu,p.Ser113Leu), four that were predicted to be benign and did not affect protein expression (p.Thr77Ile, p.Lys93Glu, p.Ala144Thr, p.Gln153Arg), three for which the results of pathogenicity predictions and western blots were contradictory (p.Arg102Trp, p.Glu138Lys, p.Gly141Asp) and two where the expression assay was inconclusive (p.Leu21Pro, p.Arg48Trp).

Wild type or variant forms of IL-36Ra were generated by transfecting HEK293 cells with myc-tagged constructs and then purifying the recombinant protein from whole cell lysates. HeLa cells which stably over-expressed the IL-36 receptor (HeLa-IL-36R) were then incubated with purified IL-36Ra, before stimulation with IL-36α. Finally, IL-8 induction was measured by ELISA (Figure 4.3). In the presence of functional IL-36Ra, IL-8 up-regulation is limited as the receptor antagonist will be competing with IL-36α for binding to IL-36R. If the antagonist is not functional, IL-8 production will be elevated.

The assay showed a clear distinction between the activity of wild type IL-36Ra and that of proteins harbouring the p.Leu27Pro or p.Ser113Leu disease alleles. In fact, the latter were associated with IL-8 levels that were comparable to those observed in the positive control, where no IL-36Ra was added (Figure 4.4). The damaging effect of the variants p.Pro21Leu, p.Arg48Trp, Pro76Leu and p.Glu138Lys was also confirmed, along with the benign nature of p.Thr77Ile and p.Gln153Arg. However, a reduced level of antagonist activity was seen with three variants which did not reduce protein expression: p.Lys93Glu, Arg102Trp and p.Ala144Thr (Table 4.6).

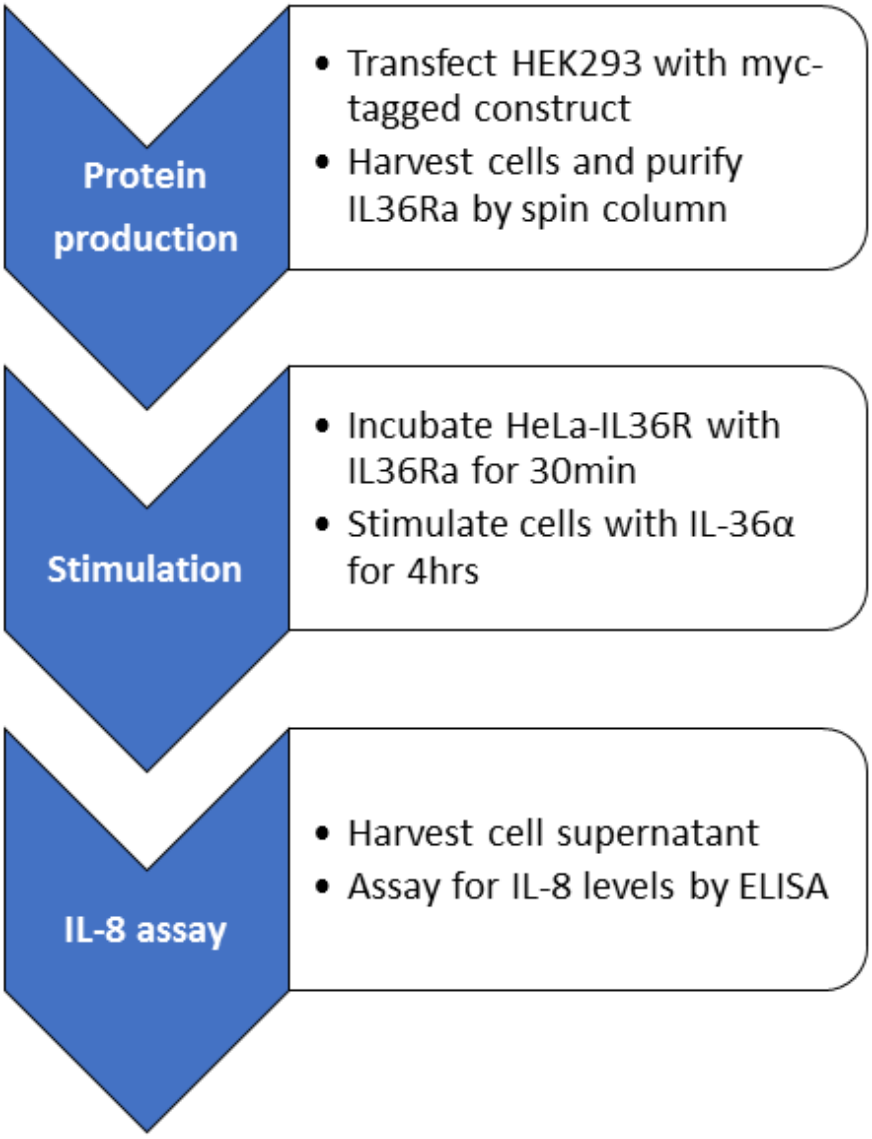


Figure 4.3 Outline of protein activity assay

Flowchart describes key steps of the procedure to assess biological activity of wild type and variant forms of IL-36Ra.

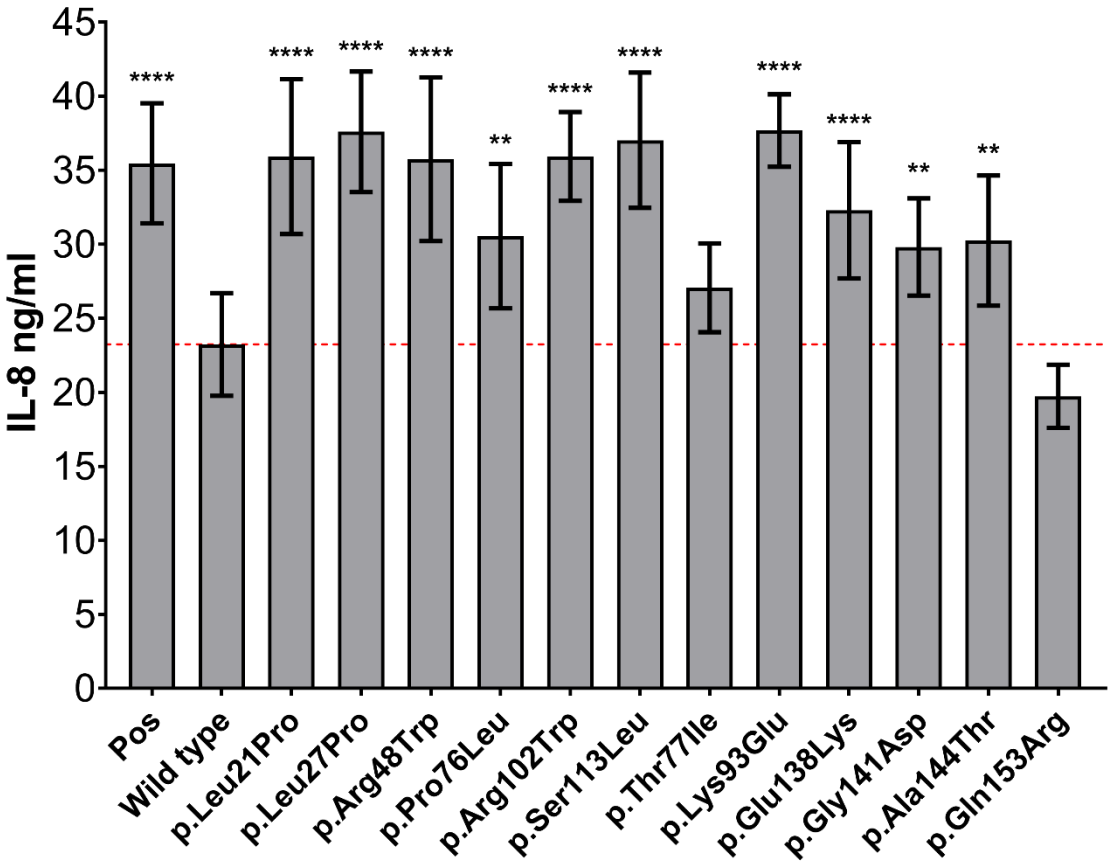


Figure 4.4 Most *IL36RN* variants affect IL-36 signal transduction

The plot shows mean IL-8 production observed in at least two stimulation experiments, each carried out with a different batch of recombinant IL-36Ra. Error bars show standard error, red line indicates mean IL-8 production in presence of wild type IL-36Ra. The mean for each variant was compared to wild type by Dunnett's multiple comparisons test. Pos: positive control (no IL-36Ra); ** $P < 0.01$; **** $P < 0.0001$

4.3 The relationship between *in-silico* predictions and the results of *in-vitro* assays

There were interesting parallels between the pathogenicity predictions and the outcome of the western blot expression assay, with consistent results obtained for 22 out of 30 alleles (Table 4.6). However, it is evident that *in-silico* tools do not perform well with the variants that disrupt the function of the protein but not the expression. For instance, p.Lys93Glu and p.Ala144Thr appeared benign on the basis of pathogenicity predictions, and did not affect protein levels, but nonetheless resulted in increased IL-8 production. The functional assay was therefore key in revealing the damaging effect of these variants. It also confirmed the pathogenicity of p.Arg102Trp, which appeared benign when analysed by western blotting (Table 4.6).

Since the activity assay was only implemented on a subset of variants, unresolved discrepancies remain between *in-silico* and *in-vitro* findings (e.g. for p.Val2Phe and p.Pro82Leu). However, as previously mentioned, in the case of p.Val2Phe the discrepancy can be explained by the detailed functional work carried out by Bal et al. This demonstrated that the effect of this variant was to prevent removal of the N-terminal methionine, which is essential for protein activity, while having no effect on protein expression (149,157).

In this context, the analysis of the 12 alleles that were fully characterised indicates that a consensus pathogenicity prediction based on the prevailing outcome from PROVEAN, MutationTaster and CADD would have 60% sensitivity and 100% specificity in differentiating *IL36RN* disease alleles from benign changes.

Table 4.6 Summary of variant assessment results

Variant	*In-silico* tools	Expression assay	Activity assay
p.Val2Phe	Damaging	Benign	
p.Leu21Pro	Damaging	Intermediate	Damaging
p.Gln25Arg	Benign	Benign	
p.Leu27Pro	Damaging	Damaging	Damaging
p.His32Arg	Benign	Benign	
p.Lys35Arg	Damaging	Intermediate	
p.Glu41Gln	Damaging	Damaging	
p.Ile42Asn	Damaging	Damaging	
p.Val44Met	Benign	Benign	
p.Asn47Ser	Benign	Benign	
p.Arg48Trp	Damaging	Intermediate	Damaging
p.Ala52Thr	Benign	Benign	
p.Cys70Arg	Damaging	Damaging	
p.Pro76Leu	Damaging	Damaging	Damaging
p.Thr77Ile	Benign	Benign	Benign
p.Pro82Leu	Damaging	Benign	
p.Lys93Glu	Benign	Benign	Damaging
p.Tyr101Phe	Benign	Benign	
p.Arg102Trp	Damaging	Benign	Damaging
p.Arg102Gln	Benign	Benign	
p.Arg103Gln	Benign	Benign	
p.Glu112Lys	Damaging	Damaging	
p.Ser113Leu	Damaging	Damaging	Damaging
p.Thr123Arg	Damaging	Damaging	
p.Thr123Met	Damaging	Damaging	
p.Glu138Lys	Benign	Damaging	Damaging
p.Gly141Asp	Benign	Damaging	Damaging
p.Ala144Thr	Benign	Benign	Damaging
p.Ile146Val	Benign	Benign	
p.Gln153Arg	Benign	Benign	Benign

4.4 Discussion

Over the past seven years, *IL36RN* variants have been assessed with a variety of *in-silico* tools and, in some cases, *in-vitro* methods. Prior to this work, the most comprehensive study was published in 2016 and included 9 of the 18 missense alleles listed in Table 4.1 (151). The expression of these variants was investigated in a similar way as has been utilised here, but protein activity levels were explored through a luciferase reporter assay. Other *in-vitro* methods used to assess the impact of changes in IL-36Ra include mass spectrometry (149), stimulation of patient cells (28,29), thermostability assays (29) and immunohistochemistry in patient skin (133,150).

These disparate means of assessment make it difficult to compare the effects of all published IL-36Ra variants on protein function. However, it is also important to note that in some instances (e.g. p.Val2Phe) a more unusual method of investigation (e.g. mass spectrometry) may be necessary to characterise the precise outcome of a missense substitution. Therefore, any future *IL36RN*-specific classifier should take into account the fact that such 'outlier' variants may not be fully described by an integration of more basic assay techniques.

Despite concerns over the reliability of pathogenicity prediction tools, this study has shown that by integrating the results from three independent tools (PROVEAN, MutationTaster and CADD) a good degree of correlation between *in-silico* and *in-vitro* findings can be achieved.

Applying the same three *in-silico* tools to all the 64 *IL36RN* missense changes reported on ExAC suggests that 60.9% (39/64) would be pathogenic. However, it is likely that the true proportion is even greater, given that the prediction tools do not perform as accurately for variants that impact protein activity but do not reduce expression. Of note, 63 of the 64 variants are rare (MAF<0.5%), which is also consistent with the notion of a damaging effect.

Therefore, it seems that the IL-36Ra protein is sensitive to missense changes, with the majority likely to reduce protein levels and/or render it inactive. This contrasts with the *IL36RN* constraint metrics that are reported on ExAC, which suggest broad tolerance to missense changes (negative z score) and nonsense alleles (probability of being loss-of-function intolerant = 0). As the coding region of *IL36RN* is only 468 nucleotides in length it seems likely that the power of the ExAC calculations has been reduced by the small size of the gene and resulting low number of possible variants (158).

It is also possible that variants in *IL36RN* are not always disadvantageous, particularly if they occur in the heterozygous state. A moderate reduction in antagonism of the IL-36 receptor, causing a small rise in the expression of IL-36 target genes, could increase the immune response within a safe and beneficial range.

Beyond these general considerations, however, it is evident that popular pathogenicity prediction tools are not sufficiently robust for use in a diagnostic setting, even when the outputs of different programs are combined in a consensus score.

As previously described, the greatest weakness of the *in-silico* approach was in assessing variants that likely affect the binding of IL-36Ra to the IL-36 receptor. Thus, the identification of the residues which are key to this interaction could improve the accuracy of pathogenicity predictions.

A recent study mapped the structure of mouse IL-36Ra onto that of human IL-36γ and found that the greatest structural differences between the two were found in the β4/5, β6/7 and β11/12 loops, which connect the β sheets found within the protein. More specifically, in mouse IL-36Ra β4/5 was shorter and β11/12 was longer. Subsequent experiments, utilising modified human IL-36Ra, indicated that loop β4/5 was particularly important for antagonist activity (31).

Examination of the residues in IL-1β that interact with IL-1RaCP (an accessory protein mediating both IL-1 and IL-36 signalling) also suggests that other key residues lie in IL-36Ra β sheets 8 and 9 (31). Interestingly, these include p.Arg103, which does not destabilise protein folding when mutated into a Gln (Table 4.6, Figure 4.5).

If the IL-36Ra/IL-36R interaction can be accurately reconstructed in 3D, it should be possible to integrate the location of key residues into an improved prediction model. Detailed examination of both the crystal structure of murine IL-36Ra (148) and the available human 'loop-swapped' models (147) may also reveal the residues most likely to disrupt the folding of the protein into its β-trefoil tertiary structure. These analyses are ongoing in the Capon lab and hold the promise of producing a reliable classifier for *IL36RN* alleles.

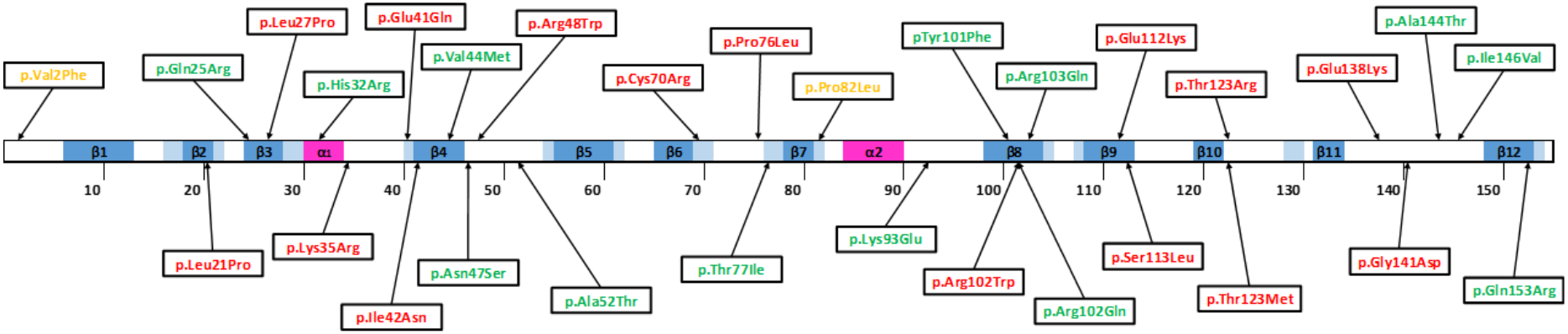


Figure 4.5 Schema of IL-36Ra showing the location of assayed variants within the protein secondary structure

Figure displays the locations of variants assayed in this study, alongside key secondary structures within IL-36Ra. Colouring of variants reflects that used in Table 4.6, with red indicating damaging, green benign and orange intermediate. Location of β-sheets (blue) and α-helices (pink) taken from Dunn et al and Gunther et al (31,32). Lighter blue highlighting indicates residues identified as being within a β-sheet by only one of these publications.

5 Identification of new candidate genes by whole exome sequencing

5.1 Analysis of eight acrodermatitis continua of Hallopeau cases

Gene discovery was initially pursued through the analysis of a tightly phenotyped dataset, comprising eight isolated ACH cases. None of the patients carried rare variants in *IL36RN*, *AP1S3* or *CARD14*; all were European and most suffered from adult-onset disease. Two affected individuals also presented with other forms of pustular psoriasis, and three had a history of plaque psoriasis (Table 5.1).

5.1.1 Preliminary data filtering

Because the cases were isolated, it was not possible to infer whether the causative variant would be likely to be homozygous or heterozygous. The data was therefore analysed twice: first assuming a recessive and then a dominant inheritance pattern (Figure 5.1).

In total the eight exomes contained 196,859 variants, an average of 24,607 changes per patient. All synonymous variants were removed, along with variants annotated as 'unknown'. Of the remaining alleles only those with an ExAC (159) global MAF ≤ 0.05 were retained (the disease itself is rare and so it was expected that the causative mutation would also occur with low-frequency). This left 15,868 variants (~1,984 per patient).

Table 5.1 Details of exome sequenced patients used in discovery analysis

Patient ID	Diagnosis	Sex	Ethnicity	Age of onset	Plaque psoriasis
S0655	ACH	F	European	51	Y
S0657	ACH	F	European	82	N
S0658	ACH	M	European	49	N
S0660	ACH	F	European	56	Y
S0661	ACH	F	European	42	N
S1948	ACH, GPP	F	European	49, 55	N
S1953	ACH, PPP	F	European	17, 17	Y
S1954	ACH	M	European	66	Y

5.1.2 Analysis of homozygous and compound heterozygous variants

If a disease allele is recessive, it will be seen in patients in the homozygous or compound heterozygous state. Starting from the set of non-synonymous variants with a MAF ≤0.05, homozygous alleles were therefore combined with potentially compound heterozygous variants for the recessive analysis. This generated a dataset of 6,316 variant calls (~790 per patient).

The data was next filtered to remove alleles with a MAF greater than 0.05 in ExAC, 1000 Genomes (72) or in-house sequencing datasets, and insertions or deletions (INDELs) of more than three bases. This step aimed to reduce the number of sequencing artefacts and left 1,928 variants. HLA alleles were also filtered out, because the coding regions have high levels of sequence similarity and so it is difficult for alignment algorithms to map exome sequencing reads to the correct gene reliably.

Any change affecting the 100 genes most frequently mutated in public exomes was also excluded, based on the work of Shyr et al., who found that rare variants in these genes are less likely to be disease-associated (120). At each filtering stage, the data set was re-examined: genes that were mutated in only one patient and heterozygous variants for which the compound partner had been removed were excluded from subsequent analysis.

In total, 462 variant calls (298 unique changes) were carried forward for assessment by pathogenicity prediction software. The exome sequencing coverage of those classified as damaging by a majority of tools (at least 3/5), and that of stop-gain substitutions was also reviewed with the Interactive Genome Viewer (IGV) (121).

At the conclusion of this analysis, no gene was found which harboured rare, damaging homozygous changes in more than one exome. One gene (*TLN2*, NM_015059) was identified where the same combination of heterozygous alleles was present in two affected individuals:

c.196G>A (chr15:g.62650143; p.Ala66Thr) and c.197C>A (chr15:g.62650144; p.Ala66Glu). However, closer examination of the sequencing reads revealed that these variants exist as a haplotype, and so could not be acting as compound heterozygotes (Figure 5.2).

Therefore, it is unlikely that the causative gene exhibits a recessive mode of inheritance, unless the genetic heterogeneity of the cohort is very substantial.

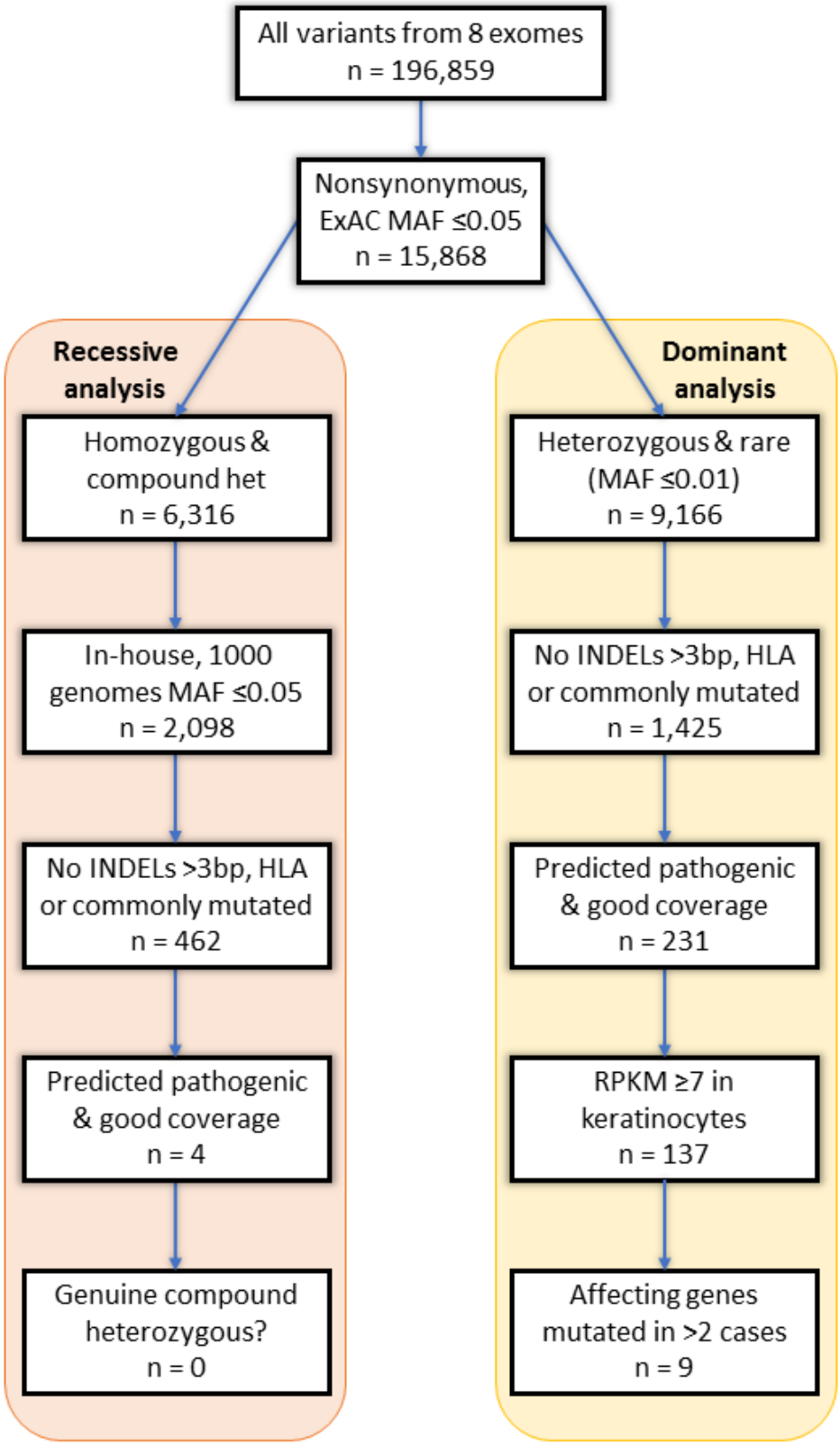


Figure 5.1 Filtering process for ACH exomes

Diagram describes the steps taken to prioritise variants seen in the 8 ACH exomes, for both recessive and dominant analysis.

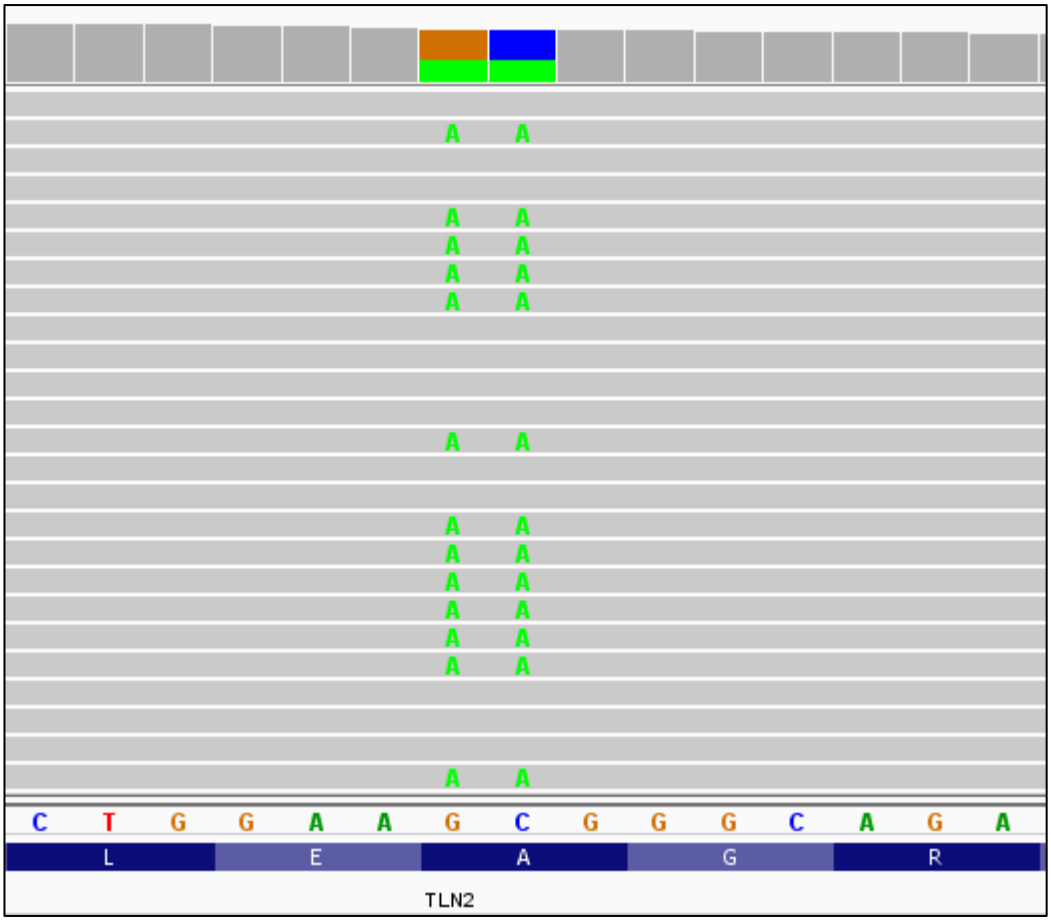


Figure 5.2 Screenshot from IGV showing the two *TLN2* variants affecting p.Ala66

The c.196G>A (p.Ala66Thr) and c.197C>A (p.Ala66Glu) changes are always seen in the same read and so are found on the same chromosome. The actual amino acid change here is therefore p.Ala66Lys.

5.1.3 Analysis of heterozygous variants

A filtering process similar to that utilised for the recessive analysis was applied when assuming dominant inheritance. The key differences were the expected zygosity of the disease allele and the lower MAF requirement (ExAC MAF of ≤0.01). Using these criteria, the original 196,859 variants were reduced to 9,166 non-synonymous heterozygous changes (1,146 per exome, on average). After examination with pathogenicity prediction tools and review of sequencing coverage, 230 variant calls in 110 genes remained.

Previous studies have shown that the known pustular psoriasis genes (*IL36RN*, *AP1S3* and *CARD14*) are all strongly expressed in keratinocytes (51,160,161). It was therefore hypothesised that this would also apply to the disease gene under investigation. RNA sequencing data, previously generated by the Capon group in primary keratinocytes (25), was examined to assess the expression pattern of the genes harbouring potentially pathogenic variants. Sixty-five loci (harbouring a total of 136 variant calls) were retained in the analysis, as they had greater than 7 reads per kilobase per million (RPKM), placing them in the top 75% of expressed genes.

The majority of the 65 genes that passed all previous stages were only mutated in two patients. However, there were three genes which showed variants in three different individuals (*ARFGAP2*, *JADE1* and *NPHP3*) (Table 5.2). These loci were examined in a further 104 pustular psoriasis exomes, previously generated by the Capon group (Table 5.3). This identified rare and potentially pathogenic alleles in all three genes. Among these, one *ARFGAP2* variant (c.1036G>A; p.Asp346Asn) was shared by two affected relatives (a Malay father and daughter, who both suffered from GPP) (Table 5.4, Figure 5.3). As the other two genes were not mutated in familial cases, *ARFGAP2* was considered the most interesting candidate gene and was selected for follow-up.

Of note, there are two translated isoforms of *ARFGAP2* (NM_032389 and NM_001242832), the shorter of which is a splicing variant lacking exon 5. However, real-time PCR undertaken in biologically relevant cell types (primary keratinocytes, HaCaT immortalised keratinocytes and peripheral blood mononuclear cells) indicated that the full-length isoform was the dominant transcript in all examined samples (Figure 5.4). Thus, no pathogenicity predictions were implemented to assess the impact of the variants on the shorter isoform.

Table 5.2 Candidate alleles observed in the original 8 patients

Gene (transcript)	Location[1] (rs ID)	Variant (cDNA)	Patients	Global MAF	Pathogenicity predictions				
					SIFT	PROVEAN	PolyPhen-2	MutationTaster	CADD[2]
ARFGAP2 (NM_032389)	chr11:47175270 (rs117324352)	p.Asn103Ser (c.308A>G)	S0655	0.0015	Damaging	Deleterious	Benign	Disease causing	21.1
	chr11:47168235 (rs373135315)	p.Val320Met (c.958G>A)	S0658	NA	Damaging	Neutral	Possibly damaging	Disease causing	25.3
	chr11:47165493 (rs142683966)	p.Gly519Ser (c.1555G>A)	S0660	0.0029	Tolerated	Deleterious	Probably damaging	Disease causing	22.8
JADE1 (NM_001287439)	chr4:128871655 (rs141808337)	p.Arg641His (c.1922G>A)	S0657 S0658	0.0087	Damaging	Neutral	Benign	Disease causing	23.1
	chr4:128872087 (rs200122043)	p.Pro785Leu (c.2354C>T)	S0665	0.0009	Damaging	Neutral	Benign	Disease causing	22.5
NPHP3 (NM_153240)	chr3:132684574 (rs34391943)	p.Ala1184Thr c.3550G>A	S0655 S0661	0.0086	Damaging	Deleterious	Probably damaging	Disease causing	24.7
	chr3:132682759 (rs143451766)	p.Ser1252Arg c.3756C>G	S0658	0.0007	Damaging	Neutral	Probably damaging	Disease causing	24.9

[1]As per GRCh38 genome build; [2]A CADD score over 15 is normally deemed pathogenic

Table 5.3 Summary demographics for the 104 exome sequenced patients in follow-up whole cohort

	Diagnosis				Sex		Ethnicity					Plaque psoriasis?		Mean age of onset (yrs)
	GPP	ACH	PPP	Multiple[1]	Male	Female	East Asian	South Asian	Malay	European	Unknown	Yes	No	
Patients (%)	19 (18%)	3 (3%)	79 (76%)	3 (3%)	20 (19%)	84 (81%)	5 (5%)	3 (3%)	8 (8%)	85 (81%)	3 (3%)	77 (74%)	19 (26%)	38.7

[1]Includes two patients with ACH + GPP and one with GPP + PPP

Table 5.4 Rare damaging variants in prioritised genes in follow-up exomes

Gene	Location[1] (rs ID)	Variant (cDNA)	Global MAF	Patient IDs	Diagnosis	Pathogenicity predictions				
						SIFT	PROVEAN	PolyPhen -2	Mutation Taster	CADD[2]
ARFGAP2 (NM_032389)	chr11:47175270 (rs117324352)	p.Asn103Ser (c.308A>G)	0.0015	S2815	PPP	Damaging	Deleterious	Benign	Disease causing	21.1
	chr11:47168157 (rs369659958)	p.Asp346Asn (c.1036G>A)	0.0002	S1057[3] S1952	GPP GPP	Tolerated	Neutral	Possibly damaging	Disease causing	27.5
	chr11:47165493 (rs142683966)	p.Gly519Ser (c.1555G>A)	0.0029	S2715 S2754	PPP PPP	Tolerated	Deleterious	Probably damaging	Disease causing	22.8
JADE1 (NM_001287439)	chr4:128871655 (rs141808337)	p.Arg641His (c.1922G>A)	0.0087	S2820	PPP	Damaging	Neutral	Benign	Disease causing	23.1

Table continued on following page

NPHP3 (NM_153240)	chr3:132684574 (rs34391943)	p.Ala1184Thr (c.3550G>A)	0.0086	S2759	PPP	Damaging	Deleterious	Probably damaging	Disease causing	24.7
	chr3:132691199 (rs201237799)	p.Gln855X (c.2563C>T)	$6.6x10^{-5}$	S2704	PPP	NA	NA	NA	Disease causing	39
	chr3:132708187 (rs141477666)	p.Arg397Cys (c.1189C>T)	0.0033	S0126 S0127 S2698 S2723	GPP GPP ACH PPP	Damaging	Neutral	Probably damaging	Disease causing	25.7
	chr3:132708219 (rs142021049)	p.Asn386Ser (c.1157A>G)	0.0017	S2777	PPP	Tolerated	Neutral	Benign	Disease causing	21.6

[1]As per GRCh38 genome build; [2]A CADD score over 15 is normally deemed pathogenic; [3]S1057 and S1952 are related (father/daughter)

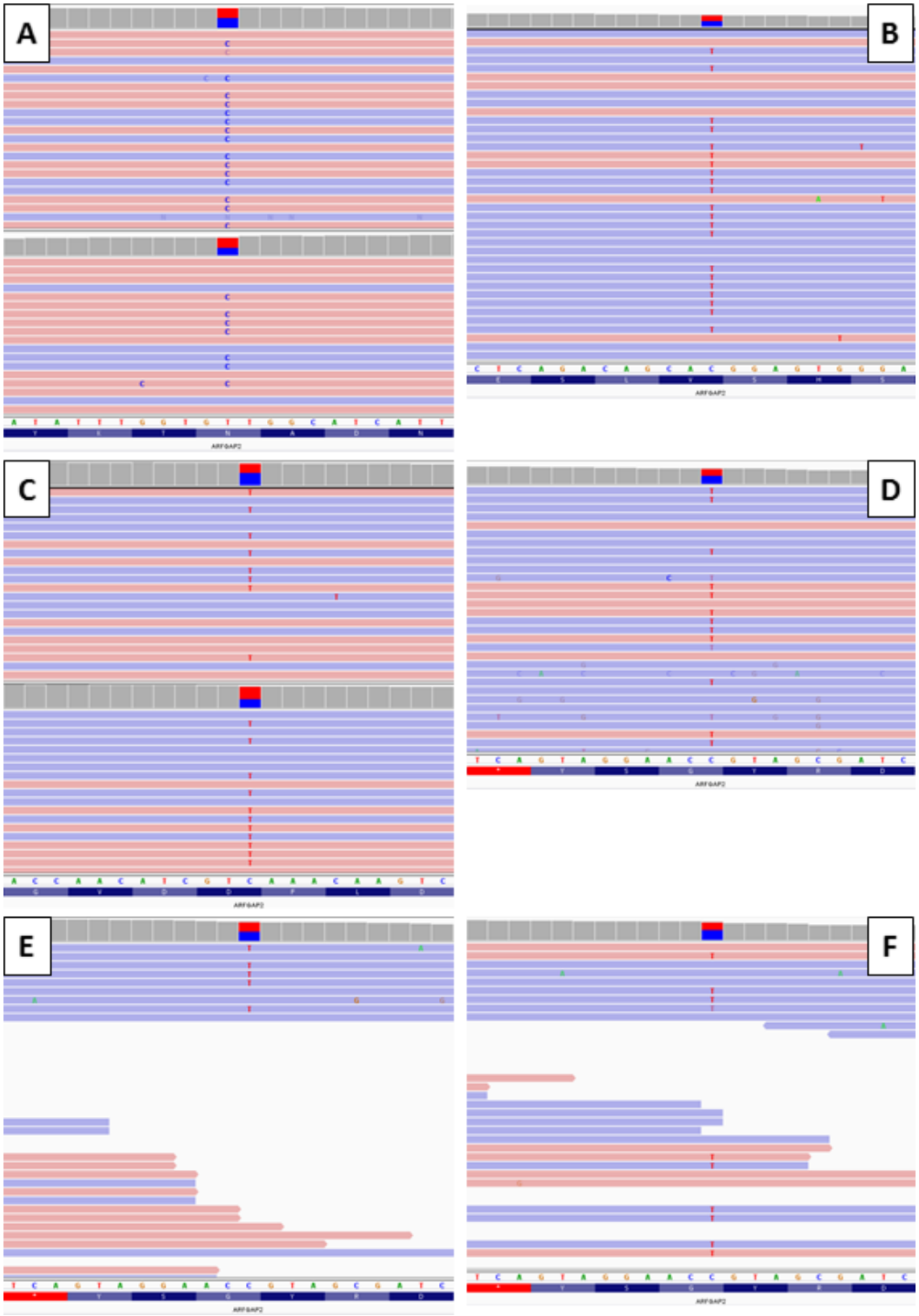


Figure 5.3 IGV visualisations of *ARFGAP2* WES variants

IGV visualisation of *ARFGAP2* variants seen in whole exome sequencing. Note that *ARFGAP2* lies on the reverse strand. **A:** c.308A>G; p.Asn103Ser. **B:** c.958G>A; p.Val320Met. **C:** c.1036G>A; p.Glu346Asp. **D-F**: c.1555G>A; p.Gly519Ser.

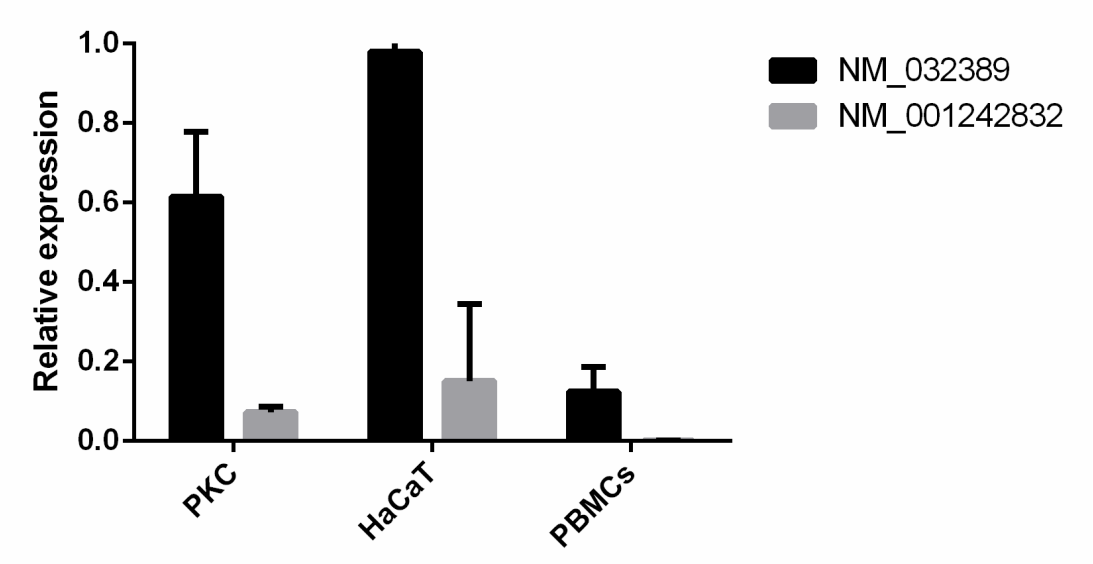


Figure 5.4 Relative expression of long (NM_032389) and short (NM_001242832) isoforms of *ARFGAP2*

Mean expression levels from two separate RNA samples are plotted with ±SD bars. The *β2M* housekeeping gene was used as an internal control. PKC: primary keratinocytes; PBMCs: peripheral blood mononuclear cells.

5.1.4 Candidate gene follow up

The *ARFGAP2* variants seen in the original eight patients and the familial case were confirmed by Sanger sequencing (Figure 5.5). Next, the 16 *ARFGAP2* coding exons were screened in an additional dataset, including 96 Asian GPP patients (Table 5.5). No rare variants were observed in the 39 affected individuals of Chinese and Indian descent. Conversely, two rare substitutions (c.674T>C; p.Leu225Pro and c.1036G>A; p.Asp346Asn) with pathogenic potential were found in three of the 57 Malay cases (Table 5.6, Figure 5.5).

Table 5.5 Summary demographics for the Asian GPP patients screened for *ARFGAP2*

	Diagnosis		Sex		Ethnicity				Plaque psoriasis?		Age of onset (yrs)		
	GPP alone	GPP & ACH	Male	Female	East Asian	South Asian	Malay	Other	Yes	No	Mean	Min	Max
Patients (%)	95 (99%)	1 (1%)	30 (31.3%)	66 (68.7%)	21 (21.9%)	16 (16.7%)	57 (59.4%)	2 (2.1%)	77 (80.2%)	19 (19.8%)	33.9	5	78

Table 5.6 Rare deleterious *ARFGAP2* variants seen in the 96 Asian GPP patients

				Pathogenicity predictions[2]				
Location[1] (rs ID)	**Variant (rs ID)**	**Patient IDs**	**Global MAF**	**SIFT**	**PROVEAN**	**PolyPhen-2**	**MutationTaster**	**CADD[3]**
chr11:47171799 (rs575856978)	p.Leu225Pro (c.674T>C)	88GPP1 106GPP1	0.0024	Damaging	Deleterious	Probably damaging	Disease causing	26.1
chr11:47168157 (rs369659958)	p.Asp346Asn[4] (c.1036G>A)	4GPP1	0.0002	Neutral	Tolerated	Possibly damaging	Disease causing	27.5

[1]As per GRCh38 genome build and transcript NM_032389; [2]Variants predicted damaging by at least 3 algorithms were considered pathogenic; [3]A CADD score over 15 is normally deemed pathogenic; [4]Also seen in exomes

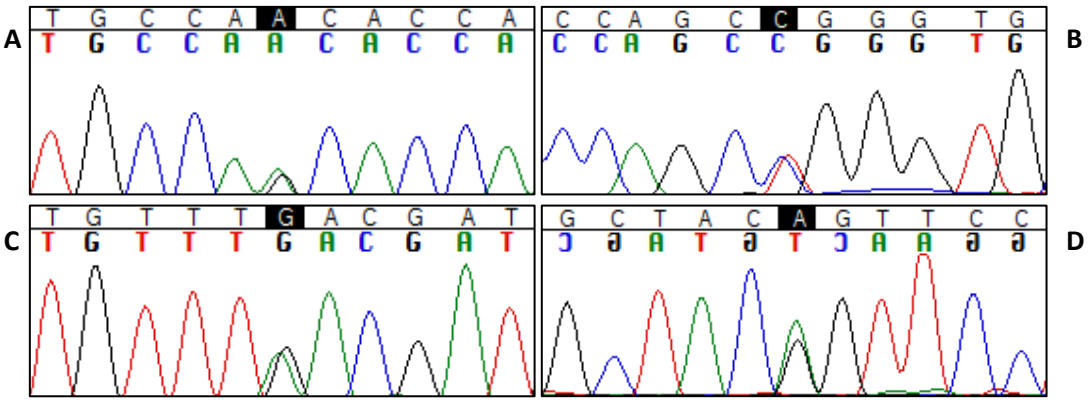


Figure 5.5 Examples of variants in *ARFGAP2* confirmed by Sanger sequencing

Chromatograms displaying heterozygous changes seen in *ARFGAP2*. **A**: c.308A>G; p.Asn103Ser. **B**: c.674T>C; p.Leu225Pro. **C**: c.1036G>A; p.Glu346Asp. **D**: c.1555G>A; Gly519Ser

5.1.5 Comparison to control populations

The association between a disease and a set of rare variants affecting a given locus can be assessed using a burden test. In its simplest form, this is implemented by comparing the combined frequency of rare alleles in cases vs controls, using a one-tailed Fisher's exact test (162).

Here, the Singapore Sequencing Malay Project (SSMP) dataset (103), which includes whole genome sequences from 96 healthy Malay individuals, was used as a control dataset for the Malay population The analysis of this cohort uncovered four individuals, each carrying a rare heterozygous and potentially damaging *ARFGAP2* change (Table 5.7). The comparison with the allele frequency measured in Malay cases returned P = 0.376 (Table 5.8), indicating that *ARFGAP2* variants are not associated with the disease in Malay patients.

For the 112 European cases (including the 8 original ACH cases and the 104 exomes in the follow-up dataset) the control population were the Non-Finnish Europeans reported in ExAC (159), where a total of 621 rare, pathogenic, alleles (135 independent variants) were seen in 32,649 individuals. Here the burden test demonstrated a significant enrichment for *ARFGAP2* variants in cases compared to controls (P = 0.009, Table 5.9). This data does not, however, reach the threshold for exome-wide significance ($P = 2.5 \times 10^{-6}$).

Table 5.7 *ARFGAP2* variants found in Malay controls

Location[1] (rs ID)	Variant (cDNA)	Allele counts[2]	Individual IDs	Pathogenicity predictions[3]				
				PROVEAN	SIFT	PolyPhen-2	CADD[4]	Mutation Taster
chr11:47171709 (rs367812497)	p.Arg255His (c.764G>A)	1/192	SSM100	Deleterious	Damaging	Benign	26.9	Disease causing
chr11:47171487 (rs374493764)	p.Asn294Tyr (c.796A>T)	1/192	SSM064	Deleterious	Damaging	Possibly damaging	26.6	Disease causing
chr11:47168235 (rs373135315)	p.Val320Met (c.958G>A)	1/192	SSM096	Neutral	Damaging	Possibly damaging	33	Disease causing
chr11:47168157 (rs369659958)	p.Asp346Asn (c.1036G>A)	1/192	SSM090	Neutral	Tolerated	Possibly damaging	25.9	Disease causing

[1]As per GRCh38 genome build and transcript NM_032389; [2]From a total of 96 individuals from the Singapore Sequencing Malay Project; [3]Variants predicted damaging by at least 3 algorithms were considered pathogenic; [4]A CADD score over 15 is normally deemed pathogenic

Table 5.8 Burden association test for *ARFGAP2* in Malay cases

Allele counts in cases[1]		**Allele counts in controls**[2]		***P* value**[3]
Alt/All	Frequency	Alt/All	Frequency	
4/124	0.0323	4/192	0.1484	0.376

[1]Familial cases counted as if a single individual; [2]Control data from SSMP; [3]One-sided *P* value from Fisher's exact test

Table 5.9 Burden association test for *ARFGAP2* in European cases

Allele counts in cases		**Allele counts in controls**[1]		***P* value**[2]
Alt/All	Frequency	Alt/All	Frequency	
6/186	0.0322	621/65298	0.0095	0.009

[1]Control data from ExAC Non-Finnish Europeans; [2]One-sided *P* value from Fisher's exact test

5.2 Analysis of a paediatric-onset generalised pustular psoriasis case

5.2.1 Case selection

A female paediatric case of generalised pustular psoriasis (disease onset at 9yrs) was whole exome sequenced prior to the onset of the study. The individual belonged to a population enriched for parental relatedness (Roma travellers) (163,164), and was the offspring of first cousins. Thus, a recessive mode of disease inheritance was assumed, where the proband was expected to have inherited two copies of the same variant, lying on identical haplotypes.

5.2.2 Filtering of the exome profile

The exome of this proband contained 24,581 variants, of which 10,441 (42.5%) were homozygous and 5,582 were also non-synonymous. Filtering to remove common changes (ExAC, 1000 Genomes or in-house MAF ≥0.05) left 101 to be investigated in more detail (Figure 5.6).

First, the haplotype background of these alleles was assessed. As it was assumed that the pathogenic variants would be identical by descent, it followed that they should map to a wider region of homozygosity. Changes were therefore retained if they lay within a homozygosity block of at least 2Mb. This was a low threshold, as stretches of homozygosity <3Mb in length are commonly found even in outbred populations (165). Following this step, 41 variants remained; 30 were single nucleotide variants (SNVs), 6 were splicing changes and 5 were insertions or deletions (INDELs).

Pathogenicity prediction software suggested that 11 SNVs could be damaging. As in the analysis of heterozygous variants in the ACH profiles, genes with low expression in keratinocytes (RPKM

<7) were removed. This left seven genes which were prioritised for further examination (Table 5.10).

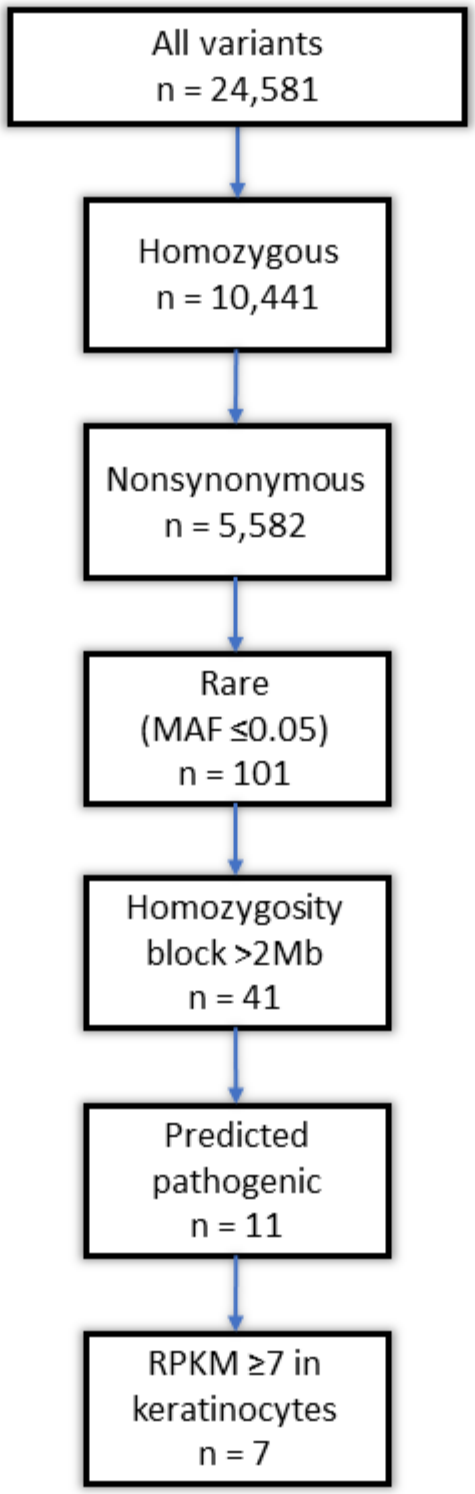


Figure 5.6 Filtering process for exome from paediatric case

The above steps were taken to prioritise variants seen in the exome of a patient suffering from paediatric-onset GPP.

Table 5.10 Candidate alleles identified in the paediatric GPP patient

Gene (Transcript)	RPKM in KCs	Location[1] (rs ID)	Variant (cDNA)	Global MAF	Pathogenicity predictions[2]				
					SIFT	PROVEAN	PolyPhen-2	Mutation Taster	CADD[3]
EVI5 (NM_001308248)	80.94	chr1:92695323 (rs143611208)	p.Ile343Thr (c.1028T>C)	0.0016	Damaging	Deleterious	Possibly damaging	Disease causing	26.7
PRTFDC1 (NM_020200)	43.24	chr10:24952573 (rs143356584)	p.Met1Ile (c.3G>A)	0.0193	Damaging	Neutral	Benign	Disease causing	22.2
SLC45A4 (NM_001286646)	20.59	chr8:141212205 (rs368838052)	p.Glu765Gln (c.2293G>C)	0.00005	Damaging	Neutral	Probably damaging	Disease causing	24.4
SPAG5 (NM_006461)	160.05	chr17:28592094 (rs143027546)	p.Val384Met (c.1150G>A)	0.0012	Damaging	Neutral	Probably damaging	Polymorphism	23.5
THEM6 (NM_016647)	52.45	chr8:142735410 (rs11540544)	p.Gly200Arg (c.598G>A)	0.0194	Tolerated	Deleterious	Benign	Disease causing	22.7
WWOX (NM_016373)	41.38	chr16:78164194 (rs369907002)	p.Ala141Thr (c.421G>A)	0.0004	Tolerated	Neutral	Probably damaging	Disease causing	27.1
ZNF33A (NM_006954)	104.77	chr10:38055200 (rs769871719)	p.Cys359Tyr (c.1076G>A)	0.00009	Damaging	Deleterious	Probably damaging	Polymorphism	22.9

[1]As per GRCh38 genome build; [2]Variants predicted damaging by 3 or more tools were considered damaging; [3]A CADD score over 15 is normally deemed pathogenic

5.2.3 Screening prioritised genes in additional exomes

To establish whether any of the prioritised genes also contained variants in other affected individuals, the 112 whole exome profiles described in the previous section (Table 5.1 and Table 5.3) were queried. Low-frequency (MAF <0.02) alleles were found in 6 of the 7 genes (Table 5.11).

The largest number of changes was found in *ZNF33A*, where 21 alleles were identified (Figure 5.7). These included a c.720G>C (p.Glu240Asp) substitution found in 11 unrelated European patients and a c.806C>T (p.Pro269Leu) allele detected in 2 Asian affected relative pairs (two siblings and two first cousins once removed). All changes were predicted to be deleterious by at least one algorithm.

The gene harbouring the second highest number of variants was *WWOX*, with 17 changes. However, one of these substitutions (c.646C>G; p.Leu103Val) was predicted to be benign by all algorithms and none of the others were shared by affected family members.

From this analysis, *ZNF33A* therefore emerged as the most promising candidate disease gene.

Table 5.11 Low-frequency variants in detected in the follow-up of candidate genes[1]

					Pathogenicity predictions				
Gene (Transcript)	**Location[2] (rs ID)**	**Variant (cDNA)**	**Patient IDs**	**Global MAF**	**SIFT**	**PROVEAN**	**PolyPhen -2**	**Mutation Taster**	**CADD[3]**
EVI5 (NM_001308248)	chr1:92513931 (rs1402834229)	p.Pro731Ser (c.2191C>T)	S2820	NA	Tolerated	Neutral	Benign	Polymorphism	21.3
	chr1:92677189 (rs146140626)	p.Thr420Met (c.1259C>T)	S2702	0.0012	Tolerated	Neutral	Probably damaging	Disease causing	23.6
	chr1:92693869 (rs200075081)	p.Asp388His (c.1162G>C)	S2752	$4.96x10^{-5}$	Damaging	Deleterious	Probably damaging	Disease causing	26.8
	chr1:92695323 (rs143611208)	p.Ile343Thr (c.1028T>C)	S0129	0.0016	Damaging	Deleterious	Possibly damaging	Disease causing	26.7
PRTFDC1 (NM_020200)	chr10:24856986 (rs145142862)	p.Gly145Arg (c.433G>A)	S2700	0.0006	Damaging	Neutral	Probably damaging	Disease causing	23.4
	chr10:24858399 (rs748761519)	p.Ile139Thr (c.416T>C)	S2647	$8.26x10^{-6}$	Damaging	Deleterious	Probably damaging	Disease causing	25.4
	chr10:24952573 (rs143356584)	p.Met1Ile (c.3G>A)	S0659, S2748	0.0193	Damaging	Neutral	Benign	Disease causing	22.2

SLC45A4 (NM_001286646)	chr8:141215862 (rs140681246)	p.Asn613Ser (c.1838A>G)	S2756, S2763, S2816	0.0064	Tolerated	Deleterious	Possibly damaging	Disease causing	16.36
	chr8:141218529 (rs148433179)	p.Glu371Lys (c.1111G>A)	S2728	0.0009	Damaging	Neutral	Probably damaging	Polymorphism	21.2
	chr8:141218840 (rs201205534)	p.Glu267Ala (c.800A>C)	S2649	$2.5x10^{-5}$	Tolerated	Neutral	Benign	Disease causing	15.09
	chr8:141219805 (rs61995884)	*p.Asn152Ser (c.455A>G)*	*S2753*	*0.0029*	*Tolerated*	*Neutral*	*Benign*	*Polymorphism*	*0.001*
SPAG5 (NM_006461)	*chr17:28578209 (rs369709868)*	*p.? (c.3429+9G>A)*	*S1057, S1952*	*$1.65x10^{-5}$*	*NA*	*NA*	*NA*	*Polymorphism*	*10.65*
	chr17:28591871 (rs144911043)	p.Arg422Trp (c.1264C>T)	S2528	0.0021	Tolerated	Neutral	Probably damaging	Polymorphism	16.5
	chr17:28592139	*p.Leu369Phe (c.1105C>T)*	*S2528*	*NA*	*Tolerated*	*Neutral*	*Benign*	*Polymorphism*	*14.88*
	chr17:28592876 (rs34977204)	*p.Gly123Asp (c.368G>A)*	*S0659, S2719, S2820*	*0.0035*	*Tolerated*	*Neutral*	*Benign*	*Polymorphism*	*4.645*

Table continued on following page

WWOX (NM_016373)	chr16:78115038 (rs144601717)	p.Pro98Leu (c.293C>T)	S2651, S2653, S2701, S2745, S2751, S2764	0.0113	Damaging	Neutral	Possibly damaging	Disease causing	20.2
	chr16:78115103 (rs141361080)	p.Arg120Trp (c.358C>T)	S2817	0.0076	Damaging	Deleterious	Probably damaging	Disease causing	26.6
	chr16:78164254 (rs779724017)	p.Ala161Thr (c.481G>A)	S2755	NA	Tolerated	Neutral	Benign	Disease causing	23
	chr16:78424910 (rs7201683)	*p.Leu216Val (c.646C>G)*	*S0657*	*0.0191*	*Tolerated*	*Neutral*	*Benign*	*Polymorphism*	*2.438*
	chr16:78425018 (rs75559202)	p.Pro252Ala (c.754C>G)	S1409, S1952, S2527	0.0232	Tolerated	Deleterious	Possibly damaging	Polymorphism	25.7
	chr16:78432512 (rs186745328)	p.Leu272Phe (c.816G>T)	S2631, S2700	0.0024	Tolerated	Neutral	Benign	Disease causing	0.002
	chr16:78432686 (rs117209694)	p.Asn330Lys (c.990C>G)	S0125	0.0002	Tolerated	Neutral	Benign	Disease causing	19.83
	chr16:79211789 (rs117065412)	p.Ser413Tyr (c.1238C>A)	S1953, S2748	0.013	Damaging	Neutral	Benign	Disease causing	15.04

Table continued on following page

ZNF33A (NM_006954)	chr10:38054478 (rs372889595)	p.Gln118His (c.354A>C)	S1414	NA	Tolerated	Neutral	Possibly damaging	Polymorphism	19.79
	chr10:38054751 (rs146079384)	p.His209Gln (c.627T>A)	S2650	0.0005	Damaging	Deleterious	Benign	Polymorphism	11.21
	chr10:38054844 (rs41276138)	p.Glu240Asp (c.720G>C)	S0129, S0661, S1058, S1953, S2628, S2631, S2652, S2698, S2708, S2752, S2815	0.0148	Tolerated	Neutral	Benign	Polymorphism	18.3
	chr10:38054930 (rs78340846)	p.Pro269Leu (c.806C>T)	S1056, S1412, S1413, S1414, S2706	0.0133	Tolerated	Neutral	Benign	Polymorphism	15.42
	chr10:38055967 (rs12256916)	p.Gly615Arg (c.1843G>A)	S2746	0.0118	Tolerated	Deleterious	Probably damaging	Polymorphism	23.1
	chr10:38056157 (rs951706084)	p.Ser678Leu (c.2033C>T)	S0127	NA	Damaging	Deleterious	Probably damaging	Polymorphism	22.1
	chr10:38056250 (rs139631793)	p.Thr709Ile (c.2126C>T)	S2707	0.0038	Tolerated	Neutral	Possibly damaging	Polymorphism	15.25

[1]Italisised rows indicate variants predicted benign by all tools; [2]As per GRCh38 genome build; [3]A CADD score greater than 15 is normally deemed pathogenic

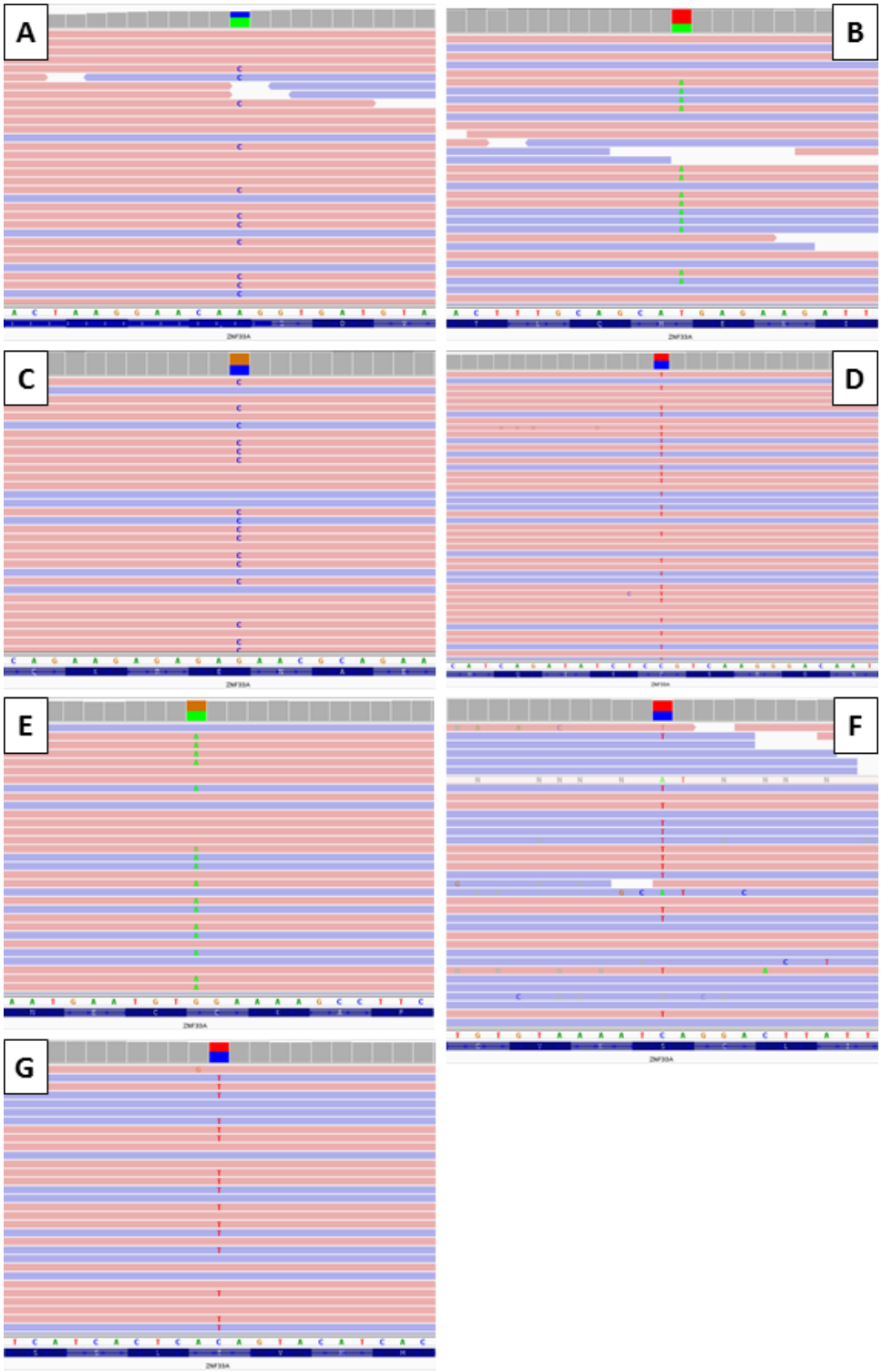


Figure 5.7 IGV visualisations of *ZNF33A* WES variants

Representative examples of IGV visualisations of *ZNF33A* variants seen in whole exome sequencing. **A:** c.354A>C; p.Gln118His. **B:** c.627T>A; p.His209Gln. **C:** c.720G>C; p.Glu240Asp. **D:** c.806C>T; p.Pro269Leu. **E:** c.1843G>A; p.Gly615Arg. **F:** c.2033C>T; p.Ser678Leu. **G:** c.2126C>T; p.Thr709Ile

5.2.4 Candidate gene screening

The original c.1076G>A (p.Cys359Tyr) change and the recurrent c.720G>C (p.Glu240Asp) and c.806C>T (p.Pro269Leu) variants were confirmed by Sanger sequencing (Figure 5.8). To further investigate the association with *ZNF33A*, the gene was next screened in an additional 372 pustular psoriasis patients (201 European, 78 Malay, 42 East Asian, 24 South Asian and 27 other) (Table 5.12). Given that p.Glu240 and p.Pro269 accounted for the majority of alleles identified in affected individuals, the screening prioritised these two variants. All cases were sequenced for the genomic segment encompassing the two changes, with a subset of 43 screened for the entire gene locus.

This analysis identified 41 additional patients carrying rare or low-frequency variants with pathogenic potential (Table 5.13). Of particular interest were the 19 European and 14 Asian patients harbouring p.Glu240Asp and p.Pro269Leu alleles respectively.

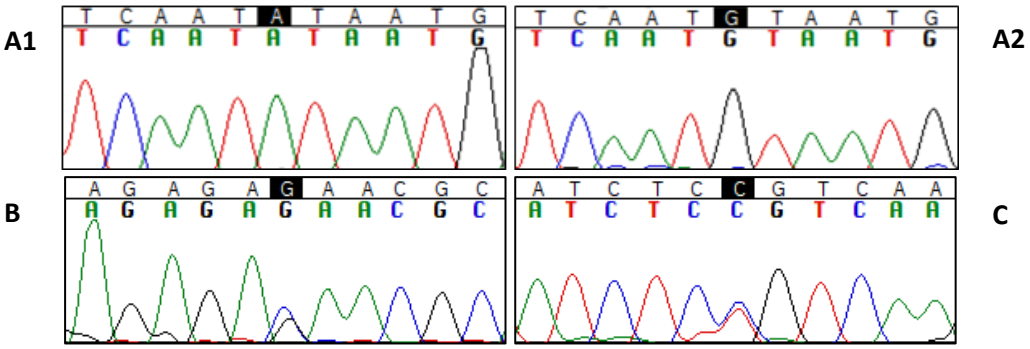


Figure 5.8 Examples of WES variants in *ZNF33A* confirmed by Sanger sequencing

Chromatograms displaying alleles seen in paediatric case and examples of recurrent variants. **A1**: Homozygous c.1076G>A (p.Cys359Tyr); **A2**: Wild-type c.1076G (p.Cys359); **B**: Heterozygous c.720G>C (p.Glu240Asp); **C**: Heterozygous c.806C>T (p.Pro269Leu)

Table 5.12 Summary demographic information for the *ZNF33A* validation cohort

	Diagnosis				Sex			Ethnicity					
	GPP	**ACH**	**PPP**	**Multiple diagnoses**	**Female**	**Male**	**Unknown**	**European**	**Malay**	**East Asian**	**South Asian**	**Other**	**Unknown**
Patients (%)	167 (44.9%)	4 (1.1%)	185 (49.7%)	16 (4.0%)	258 (69.4%)	99 (26.6%)	15 (4.0%)	201 (54.0%)	78 (21.0%)	42 (11.3%)	24 (6.5%)	6 (1.6%)	21 (5.7%)

Table 5.13 *ZNF33A* variants seen in validation cohort

				Pathogenicity predictions				
Location[1] (rs ID)	**Variant (cDNA)**	**Patient IDs**	**Global MAF**	**SIFT**	**PROVEAN**	**PolyPhen-2**	**Mutation Taster**	**CADD[2]**
chr10:38054844 (rs41276138)	p.Glu240Asp (c.720G>C)	33GPP1, 88GPP1, GBR0019, GBR0047, GBR0056, GBR0057, GBR0058, GBR0059, GBR0063, GBR0064, GBR0083, GBR0086, GBR0137, GYFAP0163, GYPLM0036, OVS0016, OVS0020, OVS0026, POPLM0018, PUS-04, PUS-08, PUS-17,PUS-19	0.0148	Tolerated	Neutral	Benign	Polymorphism	18.3
chr10:38054909 (rs200374208)	p.Leu262Pro (c.785T>C)	123GPP1	0.0025	Damaging	Deleterious	Probably damaging	Polymorphism	16.93

Table continued on following page

chr10:38054930 (rs78340846)	p.Pro269Leu (c.806C>T)	101GPP1, 127GPP1, 12GPP1, 26GPP1, 30GPP1, 3GPP1, 37GPP1, 50GPP1, 5GPP1, 67GPP1, 69GPP1, 6GPP1, 70GPP1, 76GPP1, LCPLM0002, LCPLM0005	0.0133	Tolerated	Neutral	Benign	Polymorphism	15.42
10:38054985 (rs558286901)	p.Lys287Asn (c.861G>T)	110GPP1	0.0015	Tolerated	Deleterious	Probably damaging	Polymorphism	9.648

[1]As per GRCh38 genome build and transcript NM_006954; [2]A CADD score greater than 15 is normally deemed pathogenic

5.2.5 Comparison to control populations

The frequency of the recurrent c.720G>C (p.Glu240Asp) and c.806C>T (p.Pro269Leu) variants was compared in patients and relevant control populations. A Fisher's exact test revealed that the frequency of p.Pro269Leu was not significantly increased in East Asian, Malay or South Asian patients, compared to healthy controls (*P*>0.05 in all datasets and in the meta-analysis of the three cohorts) (Table 5.14, Figure 5.9).

However, p.Glu240Asp was associated with pustular psoriasis in both the British/Irish and Non-British European populations (P = 0.006 and P = 0.018) (Table 5.15), with a meta-analysis of the two datasets yielding a *P* value of 0.46×10^{-4} (Odds ratio: 1.984, Figure 5.10).

Table 5.14 Association test for c.806C>T (p.Pro269Leu) allele in Asian patient populations

	Allele counts (%)				
	Cases	**Control**	***P* value**	**Odds ratio**	**95% CI**
East Asian	4/90 (4.4%)	962/17,226[1] (5.5%)	0.4313	0.79	0.29 – 2.1
Malay	9/169 (5.3%)	9/192[2] (4.6%)	0.484	1.14	0.44 – 2.95
South Asian	4/54 (7.4%)	1,468/30,772[3] (4.7%)	0.2314	1.60	0.58 – 4.43

[1]Data from East Asians in gnomAD; [2]Data from Singapore Sequencing Malay Project; [3]Data from South Asians in gnomAD

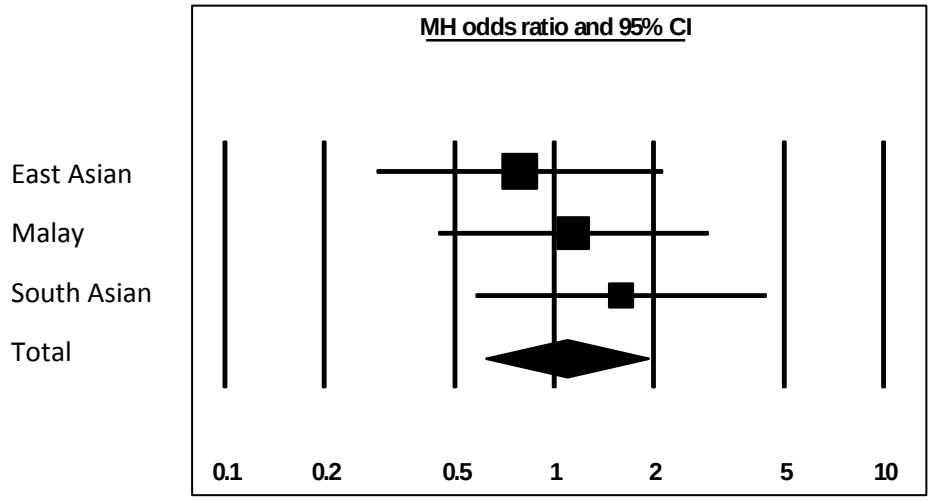


Figure 5.9 Meta-analysis of c.806C>T (p.Pro269Leu) in Asian patients

Forest plot displaying data from three Asian populations. The weights of the populations in the meta-analysis are 42.9%, 35.8% and 21.4% respectively. The squares indicate the effect size for each subgroup and the horizontal lines through the squares show the length of the confidence interval. The black diamond represents the effect size and confidence intervals when the three populations are combined. MH: Mantel-Haenszel method of meta-analysis. Generated using Comprehensive Meta-Analysis Version 3 (122).

Table 5.15 Association test for c.720G>C (p.Glu240Asp) allele in European patient populations

	Allele counts (%)				
	Cases	**Control**	***P* value**	**Odds ratio**	**95% CI**
British/ Irish	25/522 (4.8%)	254/9,428[1] (2.6%)	0.006	1.82	1.2 – 2.8
Other European	5/68 (7.4%)	2795/126,324[2] (2.2%)	0.018	3.51	1.4 – 8.7

[1]Data from UK10K consortium; [2]Data from Non-Finnish Europeans in gnomAD

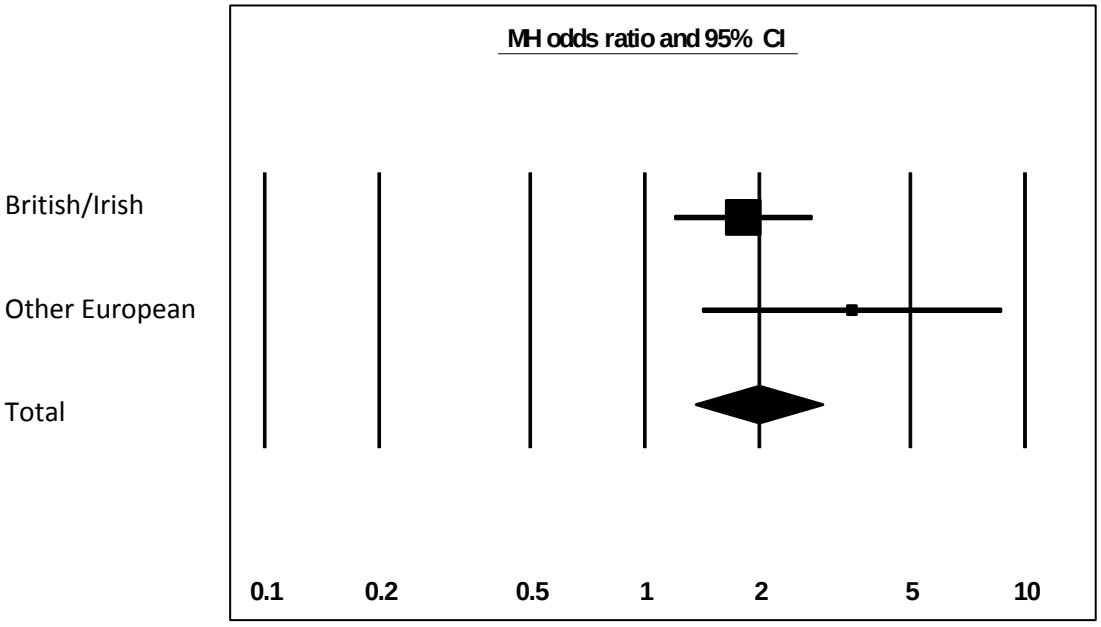


Figure 5.10 Meta-analysis of c.720G>C (p.Glu240Asp) in European patients

Forest plot displaying data from two European populations. The weight of the British/Irish population in the meta-analysis is 90.1% and for the Other European group is 9.9%. The squares indicate the effect size for each subgroup and the horizontal lines through the squares show the length of the confidence interval. The black diamond represents the effect size and confidence intervals when the two populations are combined. MH: Mantel-Haenszel method of meta-analysis. Generated using Comprehensive Meta-Analysis Version 3 (122).

5.3 Discussion

5.3.1 Analysis of acrodermatitis continua of Hallopeau cases

The attempt to identify a new disease gene through the analysis of ACH cases further highlights the genetic heterogeneity of pustular psoriasis. No gene had biallelic variants in more than one case and no gene was mutated in heterozygosity in more than three individuals. In these circumstances, power to detect a disease-associated allele was reduced and the analysis of the prioritised gene (*ARFGAP2*) yielded results that were difficult to interpret.

ARFGAP2 variants were not associated with pustular psoriasis in the Malay population, but the analysis of European cases showed an enrichment for rare pathogenic variants among patients. While this was statistically significant, exome-wide significance levels were not reached.

The result obtained in the Malay sample may reflect the extremely small size of the control cohort, which limited the power to detect a significant association. Of note, larger Malay datasets are now being generated (166). This will enable the group to investigate the role of *ARGFAP2* variants in an adequately powered sample.

While the European cohort was larger, reaching exome-wide significance would have required the screening of additional patients. For example, based on an analysis with the Genetic Power Calculator (125,167), a further 120 cases would be needed for an allele with a MAF of 0.005 and a genotype relative risk (GRR) of 5, similar to the GRR reported for *IL36RN* (44). Alternatively, the variant would need to have a larger effect (e.g. a GRR of 7).

It is also possible that the *P* value observed among European patients may represent a spurious association, with the genuine disease allele erroneously removed during the filtering process. In this context, relaxing the MAF or pathogenicity cut-offs may reveal additional candidate genes

for follow-up. It should also be remembered that the use of WES reduces the likelihood of detecting large duplications or deletions, which are better identified by whole genome sequencing. Finally, although it is certainly plausible that the causative variants could demonstrate a digenic pattern of inheritance, a digenic analysis was not carried out here. With data from only 8 patients, it is likely that the study would have lacked power. Additionally, an extremely stringent significance threshold would have been necessary to compensate for the degree of multiple testing present in this type of analysis.

5.3.1.1 The function of ARFGAP2

While it is as yet unclear whether the association between *ARFGAP2* alleles and pustular psoriasis is genuine, it is interesting to note the overlap in function with a known disease associated protein, APσ1C, a subunit of the AP-1 complex. This is recruited by Arf1-GTP during the formation of a clathrin-coated vesicle. Before the vesicle can fuse with its target membrane, ARf1 must hydrolyse GTP to GDP in order for the coat to dissociate. ADP Ribosylation Factor GTPase Activating Protein 2 (ARFGAP2) is required for the hydrolysis of GTP by Arf1, which has very low intrinsic GTPase activity (168). Therefore, if the function of ARFGAP2 is disrupted, so to would be aspects of AP-1 dependent intracellular trafficking, providing a possible role for ARFGAP2 in disease aetiology.

5.3.2 Analysis of paediatric case

5.3.2.1 Association with pustular psoriasis

Stepwise filtering of a whole exome profile identified a rare and damaging homozygous variant (c.1076G>A; p.Cys359Tyr), lying within *ZNF33A*. This gene is highly expressed in a disease relevant cell type and was located in a large region of homozygosity (25.8Mb).

Examination of additional exome profiles revealed two recurring low frequency variants (c.720G>C; p.Glu240Asp and c.806C>T; p.Pro269Leu), which were found in a number of unrelated individuals. Although these changes were only observed in the heterozygous state, the results obtained in the *IL36RN* analysis show that a disease locus can harbour bi-allelic variants in some patients and mono-allelic defects in others (44).

Further patient screening demonstrated that c.720G>C (p.Glu240Asp) is associated with pustular psoriasis at a level approaching exome wide significance. The study has therefore identified a set of disease alleles with varying effect sizes. These include an extremely rare change with strong indicators of pathogenicity (p.Cys359Tyr) and a low frequency variant (p.Glu240Asp) with a more subtle effect on protein function, indicated by a moderately sized odds ratio.

In this context, it is important to note that the pathogenicity predictions for c.720G>C (p.Glu240Asp) are weak, as the damaging effect of the variant is supported by a single algorithm (CADD). However, pathogenicity predictions are likely to be less reliable for recently evolved genes. And indeed, this is the case for *ZNF33A*, which only has orthologues in primates (106,169). Of interest, Itan et al demonstrated that the correct CADD, SIFT and PolyPhen-2 thresholds for pathogenicity vary on a gene-to-gene basis, indicating that, where possible, a gene-specific approach to variant assessment is more appropriate (99).

While the recent evolution of *ZNF33A*, and the resulting impact on the effectiveness of pathogenicity tools, may explain the weaker predictions, there is a heightened need for *in-vitro* confirmation of a damaging effect caused by the c.720G>C (p.Glu240Asp) allele. Mutagenised constructs are being generated so that the expression, stability and intracellular localisation of wild-type and variant proteins can be compared.

5.3.2.2 The function of ZNF33A

ZNF33A encodes a zinc-finger transcription factor of the KRAB (Krüppel-associated box) domain-containing family. These proteins bind DNA via multiple C2H2 zinc-finger motifs, while their KRAB domain associated with the universal repressor known as KRAB Associated Protein 1 (KAP1) (170). Thus, KRAB zinc fingers confer target specificity to the silencing activity of KAP1.

The consensus sequence for the C2H2 motif is ϕ-X-Cys-X(2-4)-Cys-X3-ϕ-X5-ϕX2-His-X(3,4)-His (where ϕ is a hydrophobic residue and X can be any residue), with the key cysteine and histidine residues binding a zinc ion, which stabilises the structure of the finger (170). *ZNF33A* encompasses 16 C2H2 motifs, encoded by exon 5. Notably, the homozygous p.Cys359Tyr change, identified in the original paediatric case, affects the first key cysteine residue within the second C2H2 motif. It is therefore likely to affect formation of this zinc finger structure. Three additional changes (p.Gly615Arg, p.Ser678Leu and p.Thr709Ile) also map to C2H2 motifs, altering broadly conserved residues. The remaining variants are mostly found in the linker region between the KRAB domain and the C2H2 zinc fingers, where they may affect protein folding and stability (Figure 5.11).

There are over 700 KRAB-ZNFs within the human genome, with a wide range of transcriptional targets (171). While the specific function of ZNF33A is currently unknown, it is noteworthy that other KRAB proteins have been involved in various aspects of immune function, most notably antiviral defence and NF-κB signalling. Thus, it is tempting to speculate that loss-of-function *ZNF33A* alleles may cause excessive inflammatory signalling.

Efforts are ongoing within the Capon group to elucidate the mechanisms whereby *ZNF33A* contributes to disease pathogenesis. As the protein binds DNA, it would be particularly informative to identify its targets. Of note, a systematic ChIP-Seq study of 39 KRAB proteins (including ZNF33A) has been published and the raw data is publicly available (172). The analysis

of this dataset may provide insights into the genes that are regulated by *ZNF33A* and the transcriptional networks that are altered by pustular psoriasis disease alleles.

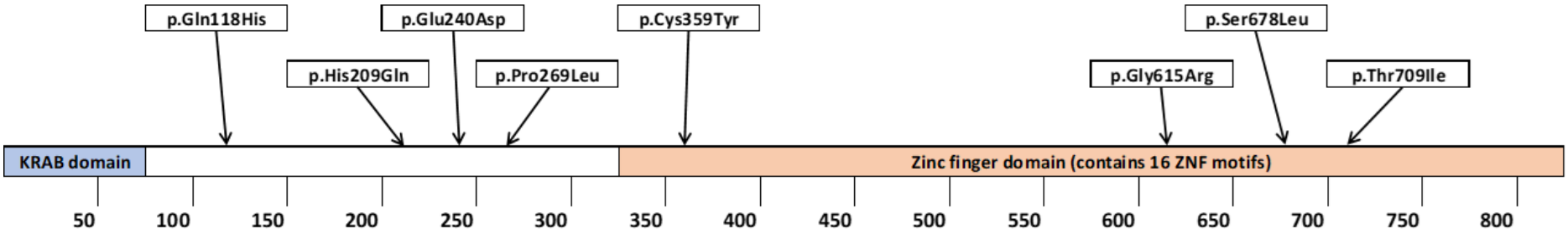


Figure 5.11 Schematic of ZNF33A variants seen in exomes

Four of the eight variants identified in patients map to C2H2 motifs within the zinc finger domain, with p.Cys359Tyr affecting a key residue within the second of sixteen motifs. The remaining four variants lie within a region without annotated function but may affect the overall folding of the protein. Location of domains taken from Uniprot (173,174).

6 Final discussion

6.1 Genetic advances

A pustular form of psoriasis was first described in 1910 (175), but over 100 years later our understanding of disease aetiology is still incomplete, and available treatments are often ineffective (8). This contrasts with plaque psoriasis, where a more extensive knowledge of the immunopathogenesis has transformed treatment efficacy, particularly for severe disease (22,176).

The first genetic determinants of plaque psoriasis were identified in the 1990s (177), whereas *IL36RN* alleles were only uncovered in 2011 (28,29). Work reported here (section 3.3.2) and published by others has shown that a small number of recurrent and well-characterised variants (p.Leu27Pro, p.Arg48Trp, p.Ser113Leu, c.115+6T>C) account for the majority of *IL36RN* disease alleles in northern Europe, north Africa and East Asia (2,15,28,29).

However, as *IL36RN* screening is extended to further populations, the incidence of private variants will increase. Of note, *IL36RN* is now included on diagnostic gene panels for autoinflammatory syndromes (85), so new sequence changes will likely be identified in patients without a pre-existing GPP diagnosis. This will further complicate the interpretation of genetic findings.

It is therefore clear that reliable *in-silico* predictions will be needed for *IL36RN* alleles which have not been functionally characterised. This study has demonstrated that currently available algorithms have their limitations, particularly in the case of variants that do not alter protein expression. However, databases such as ClinVar (178) or Infevers (179) may provide important

guidance – a rare sequence change that has already been seen in an individual with a similar phenotype is more likely to be pathogenic.

Following the identification of *IL36RN* variants, disease-associated alleles have been identified in two additional genes (*AP1S3* and *CARD14*) (46,48). However, the majority of cases remain unaccounted for (41).

This study has demonstrated that genetic heterogeneity is present even in cohorts of robustly phenotyped patients. One notable example is that Malay subjects had much higher rates of plaque psoriasis. It would therefore be interesting to explore whether this can be linked to a variant that is both Malay specific and associated with plaque psoriasis.

While *IL36RN* and *AP1S3* alleles accounted for a sizeable minority of generalised pustular psoriasis (GPP) and acrodermatitis of Hallopeau (ACH) cases, their frequency in PPP was significantly lower. It therefore seems likely that the genetic architecture of PPP is distinct from that of GPP, a hypothesis which is reflected in contrasting demographic and phenotypic data.

Here, genetic heterogeneity has also affected attempts to identify further disease genes by whole exome sequencing (WES). While a promising *ZNF33A* change (c.1076G>A; p.Cys359Tyr) was uncovered in a GPP subject who was the offspring of first cousins, the *ZNF33A* variants found in the broader patient cohort are relatively common in the general population (minor allele frequency >1%) and seem to have a modest impact on protein function. As such, these *ZNF33A* alleles are more likely to act as modifiers than major genetic determinants of the disease.

Therefore, it appears increasingly probable that the genetics of pustular psoriasis is more complex than initially thought and that the 'low-hanging fruits' have already been discovered. In this context, it seems likely that future advances will come from alternative investigative

routes. If further analyses of WES data do take place, it would be interesting to utilise the recently published technique of 'blacklisting' variants seen frequently in private exome cohorts but not public databases, which has been shown to be effective at removing non-pathogenic variants (80).

In pulmonary arterial hypertension (PAH), a rare condition with an autoinflammatory component, epigenetic dysregulation has been found to play a role in disease development (180). It is possible that changes affecting non-coding RNAs, histone modifications and DNA methylation may also underlie some cases of pustular psoriasis, especially given the marked association between palmoplantar pustulosis and smoking.

In amyotrophic lateral sclerosis (ALS) which, like PPP, is rare and genetically heterogeneous, a disease-associated locus (*C21orf2*) has been identified by undertaking a genome-wide association scan (GWAS) (181,182).

Classical GWAS, which are intended to detect common variants with small effect sizes, typically require several thousand cases. However, if the underlying disease alleles have a moderate effect then smaller datasets can still yield positive results. In fact, a new PAH gene was recently identified through the analysis of only 256 affected individuals (183).

As the underlying genetic architecture could be similar in palmoplantar pustulosis, a GWAS (or exome-wide association study (EWAS)) may be feasible through the involvement of an international consortium such as ERASPEN. Patients are being actively ascertained to this project, as well as the PLUM study, and so the size of patient cohorts continues to grow.

By utilising the common CRF described in this study, it may also be possible to stratify affected individuals into more homogenous groups. This would be of particular interest in palmoplantar

pustulosis, where in-depth phenotyping could identify further correlations between genotypic and clinical features, leading to an improved understanding of the disease.

6.2 Interaction with environmental factors

It is clear that pustular psoriasis develops as a result of both genetic and environmental factors, not least because of the wide range of phenotypes that can be observed in individuals who carry the same disease allele. One way to explain this is that genetic disrupt the balance between pro- and anti-inflammatory factors, such as the IL-36 receptor agonists and antagonist, or the levels of NF-κB mediated gene expression. However, these deviations are generally small enough that individuals remain apparently unaffected until they are exposed to an environmental trigger which further tips the balance. Following the same logic, the more damaging the alleles carried, the smaller the environmental trigger need be. In the mouse model with complete knock-out of IL-36Ra, pups generally die young with significant skin abnormalities (38). In humans the effect of carrying homozygous disease alleles is not so severe, but is significantly associated with earlier onset (15).

As described in section 1.1.4, the IL-36 pathway can drive pro-inflammatory feedback loops that lead to increased neutrophil infiltration, pustule formation and, in some GPP cases, systemic involvement. The age at which a trigger is encountered, or the degree to which one is exposed, could then help to explain the range of clinical presentations that are seen.

One aspect of pustular psoriasis pathogenesis that remains poorly understood is the systemic nature of some disease triggers, such as upper respiratory tract infections, stress or vaccination. Presumably some individuals possess a genetic or epigenetic background that leaves them predisposed to systemic effects; if this could be better understood then it would allow for identification of patients at greater risk of extracutaneous involvement. Ongoing work carried

out in the Capon group indicates that IL-36 has systemic effects on plasmacytoid dendritic cells (pDC), which may amplify the response to infections.

Women are more likely to develop pustular psoriasis than men and disease onset is generally pre/peri-menopausal (mean age of onset for women within the cohort described here was 41.2yrs). In addition, cases of pustular psoriasis during pregnancy usually resolve after delivery (6). The implication is therefore that changes in female hormones may be one of the more common disease triggers. The finding that almost all subjects who carried an *AP1S3* variant were female (section 3.3.3) may reveal a route into understanding the female bias seen in pustular psoriasis.

6.3 Translational implications

The most recent and exciting development in pustular psoriasis research has been the emergence of a new class of biologics that act by blocking IL-36 signal transduction. The potential of this therapeutic approach was initially shown by Mahil et al (25), who reported that IL-36 receptor deficiency does not adversely affect human immune function. This observation led the authors to suggest that pharmaceutical blockade of the same receptor would be, in all likelihood, safely tolerated.

This prediction was borne out by the results of an open label, phase I trial (NCT02978690), conducted by Boehringer Ingelheim in generalised pustular psoriasis patients. Seven affected individuals (flaring at the time of treatment) received a single dose of BI 655130, an anti-IL-36 receptor antibody. The primary endpoint of safety and tolerability was successfully met, with no severe adverse events or infections reported. In addition, all seven patients achieved a clear or nearly clear status within four weeks, with five maintaining it until the conclusion of the trial (20 weeks) (184,185). Notably, only three of the seven patients carried an *IL36RN* variant. This

reinforces the notion, also supported by transcriptomic studies (25,56), that IL-36 is a key driver in GPP, regardless of the patient variant status.

A second IL-36 receptor neutralising antibody, ANB019, has been produced by AnaptysBio (27) and will soon be tested in a multi-centre trial, led by St John's Institute of Dermatology (NCT03619902). This will involve 10 affected individuals and will include genetic and transcriptomic studies of patient material.

Both Boehringer Ingelheim and AnaptysBio plan to further test their molecules in PPP patients (NCT03135548 and NCT03633396 studies, respectively). However, IL-36 blockade may be less beneficial in these individuals, due to the differences between GPP and PPP. In fact, the data obtained in this study indicate that *IL36RN* alleles are found in less than 5% of PPP cases.

The Capon group have also exome sequenced 100 unrelated PPP patients and have seen very little evidence for involvement of genes related to IL-36 signal transduction. Further transcriptomic studies are needed to complement these genetic observations and clarify to what extent IL-36 de-regulation is a feature of PPP.

6.4 Conclusions

The future for pustular psoriasis patients seems brighter than just a decade ago. In particular, significant progress has been made in the understanding of GPP. This demonstrates the translational potential of genetic studies and how a single breakthrough can, within just a few years, lead to exciting new drugs entering clinical trials.

Although more frequent, PPP remains poorly understood at the genetic level. Whole exome data generated by the Capon group suggest that this form of pustular psoriasis is particularly heterogeneous. The ascertainment of large patient samples through multi-centre studies may

help overcome this barrier. Thus, the progress that has been made in forming international collaborations holds the promise of facilitating the identification of new disease genes and promoting a better understanding of PPP.

Online resources

CADD	https://cadd.gs.washington.edu/
dbNSFP	https://sites.google.com/site/jpopgen/dbNSFP
Ensembl	http://www.ensembl.org/index.html
ERASPEN	http://eraspen.eu/
ExAC	http://exac.broadinstitute.org/
Genetic Power Calculator	http://zzz.bwh.harvard.edu/gpc/cc2.html
gnomAD	http://gnomad.broadinstitute.org
Human Splicing Finder	http://www.umd.be/HSF3/
Infevers	https://infevers.umai-montpellier.fr/web/index.php
QuikChange Primer Design	https://www.chem.agilent.com/store/primerDesignProgram.jsp
MutationTaster	http://www.mutationtaster.org/
PolyPhen-2	http://genetics.bwh.harvard.edu/pph2/bgi.shtml
Primer-BLAST	https://www.ncbi.nlm.nih.gov/tools/primer-blast/
Primer3	http://primer3.ut.ee
Protein extinction coefficient calculator	http://www.biomol.net/en/tools/proteinextinction.htm
PROVEAN + SIFT	http://provean.jcvi.org/genome_submit_2.php?species=human
Spliceman	http://fairbrother.biomed.brown.edu/spliceman/
T-Coffee	http://tcoffee.crg.cat/
UK10K	https://www.uk10k.org/
UniProt	https://www.uniprot.org/
wANNOVAR	http://wannovar.wglab.org/
WebLogo3	http://weblogo.threeplusone.com

Bibliography

1. Navarini AA, Burden AD, Capon F, Mrowietz U, Puig L, Köks S, et al. European consensus statement on phenotypes of pustular psoriasis. *Journal of the European Academy of Dermatology and Venereology*. 2017;31(11): 1792–1799. Available from: doi:10.1111/jdv.14386

2. Sugiura K. The genetic background of generalized pustular psoriasis: IL36RN mutations and CARD14 gain-of-function variants. *Journal of Dermatological Science*. 2014;74(3): 187–192. Available from: doi:10.1016/j.jdermsci.2014.02.006

3. Raychaudhuri SK, Maverakis E, Raychaudhuri SP. Diagnosis and classification of psoriasis. *Autoimmunity Reviews*. 2014;13(4–5): 490–495. Available from: doi:10.1016/j.autrev.2014.01.008

4. Augey F, Renaudier P, Nicolas JF. Generalized pustular psoriasis (Zumbusch): A French epidemiological survey. *European Journal of Dermatology*. 2006;16(6): 669–673. Available from: doi:10.1684/ejd.2006.0003

5. World Health Organisation. *Global report on psoriasis*. [Online] 2016. Available from: http://www.who.int/iris/handle/10665/204417

6. Hoegler KM, John AM, Handler MZ, Schwartz RA. Generalized pustular psoriasis: a review and update on treatment. *Journal of the European Academy of Dermatology and Venereology*. 2018;12: 1–25. Available from: doi:10.1111/jdv.14949

7. Viguier M, Allez M, Zagdanski AM, Bertheau P, De Kerviler E, Rybojad M, et al. High frequency of cholestasis in generalized pustular psoriasis: Evidence for neutrophilic involvement of the biliary tract. *Hepatology*. 2004;40(2): 452–458. Available from:

doi:10.1002/hep.20305

8. Bachelez H. Pustular psoriasis and related pustular skin diseases. *British Journal of Dermatology*. 2018;178(3): 614–618. Available from: doi:10.1111/bjd.16232

9. Boehner A, Navarini AA, Eyerich K. Generalized pustular psoriasis - A model disease for specific targeted immunotherapy, systematic review. *Experimental Dermatology*. 2018;(3): 1–9. Available from: doi:10.1111/exd.13699

10. De Oliveira ST, Maragno L, Arnone M, Takahashi MDF, Romiti R. Generalized pustular psoriasis in childhood. *Pediatric Dermatology*. 2010;27(4): 349–354. Available from: doi:10.1111/j.1525-1470.2010.01084.x

11. Borges-Costa J, Silva R, Gonçalves L, Filipe P, Soares De Almeida L, Gomes MM. Clinical and laboratory features in acute generalized pustular psoriasis: A retrospective study of 34 patients. *American Journal of Clinical Dermatology*. 2011;12(4): 271–276. Available from: doi:10.2165/11586900-000000000-00000

12. Ohkawara A, Yasuda H, Kobayashi H, Inaba Y, Ogawa H, Hashimoto I, et al. Generalized pustular psoriasis in Japan: Two distinct groups formed by differences in symptoms and genetic background. *Acta Dermato-Venereologica*. 1996;76(1): 68–71. Available from: http://www.ncbi.nlm.nih.gov/pubmed/8721499

13. Baker H, Ryan TJ. Generalized pustular psoriasis. A clinical and epidemiological study of 104 cases. *British Journal of Dermatology*. 1968;80(12): 771–793. Available from: doi:10.1111/j.1365-2133.1968.tb11947.x

14. Iizuka H, Takahashi H, Ishida-Yamamoto A. Pathophysiology of generalized pustular psoriasis. *Archives of Dermatological Research*. 2003;295 Suppl(S1): S55-9. Available

from: doi:10.1007/s00403-002-0372-5

15. Hussain S, Berki DM, Choon S-E, Burden a. D, Allen MH, Arostegui JI, et al. IL36RN mutations define a severe autoinflammatory phenotype of generalized pustular psoriasis. *Journal of Allergy and Clinical Immunology*. 2014;135(4): 1067-70.e9. Available from: doi:10.1016/j.jaci.2014.09.043

16. Alvarado SA, Muñoz-Mendoza D, Bahna SL. High-risk drug rashes. *Annals of Allergy, Asthma & Immunology*. 2018; Available from: doi:10.1016/J.ANAI.2018.05.022

17. Mansouri B, Benjegerdes K, Hyde K, Kivelevitch D. Pustular psoriasis: pathophysiology and current treatment perspectives. *Psoriasis: Targets and Therapy*. 2016;6: 131–144. Available from: doi:10.2147/PTT.S98954

18. Kubota K, Kamijima Y, Sato T, Ooba N, Koide D, Iizuka H, et al. Epidemiology of psoriasis and palmoplantar pustulosis: A nationwide study using the Japanese national claims database. *BMJ Open*. 2015;5(1): e006450. Available from: doi:10.1136/bmjopen-2014-006450

19. Hellgren L, Mobacken H. Pustulosis palmaris et plantaris. Prevalence, clinical observations and prognosis. *Acta Dermato-Venereologica*. 1971;51(4): 284–288. Available from: http://www.ncbi.nlm.nih.gov/pubmed/4105777

20. De Waal AC, Van De Kerkhof PCM. Pustulosis palmoplantaris is a disease distinct from psoriasis. *Journal of Dermatological Treatment*. 2011;22(2): 102–105. Available from: doi:10.3109/09546631003636817

21. Misiak-Galazka M, Wolska H, Rudnicka L. What do we know about palmoplantar pustulosis? *Journal of the European Academy of Dermatology and Venereology*.

2017;31(1): 38–44. Available from: doi:10.1111/jdv.13846

22. Griffiths CE, Barker JN. Pathogenesis and clinical features of psoriasis. *Lancet*. 2007;370(9583): 263–271. Available from: doi:10.1016/S0140-6736(07)61128-3

23. Robinson A, Van Voorhees AS, Hsu S, Korman NJ, Lebwohl MG, Bebo BF, et al. Treatment of pustular psoriasis: From the medical board of the National Psoriasis Foundation. *Journal of the American Academy of Dermatology*. 2012;67(3): 459–477. Available from: doi:10.1016/j.jaad.2011.01.032

24. Mrowietz U, Bachelez H, Burden AD, Rissler M, Sieder C, Orsenigo R, et al. Secukinumab for moderate-to-severe palmoplantar pustular psoriasis: Results of the 2PRECISE study. *Journal of the American Academy of Dermatology*. 2019; Available from: doi:10.1016/j.jaad.2019.01.066

25. Mahil SK, Catapano M, Di Meglio P, Dand N, Ahlfors H, Carr IM, et al. An analysis of IL-36 signature genes and individuals with IL1RL2 knockout mutations validates IL-36 as a psoriasis therapeutic target. *Science Translational Medicine*. 2017;9(411): eaan2514. Available from: doi:10.1126/scitranslmed.aan2514

26. *Boehringer Ingelheim R&D pushes to transcend disease boundaries*. [Online] Available from: https://www.boehringer-ingelheim.com/press-release/transcending-disease-boundaries

27. Khanskaya I, Pinkstaff J, Marino MH, Savall T, Li J, Londei M. *A Phase 1 Study of ANB019, an Anti-Interleukin-36-Receptor (IL-36R) Monoclonal Antibody, in Healthy Volunteers*. [Online] Available from: https://www2.anaptysbio.com/wp-content/uploads/ANB019-Phase-1-Study-Poster-EAACI-2018.pdf

28. Onoufriadis A, Simpson MA, Pink AE, Di Meglio P, Smith CH, Pullabhatla V, et al. Mutations in IL36RN/IL1F5 are associated with the severe episodic inflammatory skin disease known as generalized pustular psoriasis. *The American Journal of Human Genetics*. 2011;89(3): 432–437. Available from: doi:10.1016/j.ajhg.2011.07.022

29. Marrakchi S, Guigue P, Renshaw BR, Puel A, Pei X-Y, Fraitag S, et al. Interleukin-36-receptor antagonist deficiency and generalized pustular psoriasis. *The New England Journal of Medicine*. 2011;365(7): 620–628. Available from: doi:10.1056/NEJMoa1013068

30. Gresnigt MS, Van de Veerdonk FL. Biology of IL-36 cytokines and their role in disease. *Seminars in Immunology*. 2013;25(6): 458–465. Available from: doi:10.1016/j.smim.2013.11.003

31. Günther S, Sundberg EJ. Molecular Determinants of Agonist and Antagonist Signaling through the IL-36 Receptor. *The Journal of Immunology*. 2014;193(2): 921–930. Available from: doi:10.4049/jimmunol.1400538

32. Dunn EF, Gay NJ, Bristow AF, Gearing DP, O'Neill LAJ, Pei XY. High-resolution structure of murine interleukin 1 homologue IL-1F5 reveals unique loop conformations for receptor binding specificity. *Biochemistry*. 2003;42(37): 10938–10944. Available from: doi:10.1021/bi0341197

33. Gabay C, Towne JE. Regulation and function of interleukin-36 cytokines in homeostasis and pathological conditions. *Journal of Leukocyte Biology*. 2015;97(4): 645–652. Available from: doi:10.1189/jlb.3RI1014-495R

34. Henry CM, Sullivan GP, Clancy DM, Afonina IS, Kulms D, Martin SJ. Neutrophil-Derived Proteases Escalate Inflammation through Activation of IL-36 Family Cytokines. *Cell*

Reports. 2016;14(4): 708–722. Available from: doi:10.1016/j.celrep.2015.12.072

35. Yi G, Ybe JA, Saha SS, Caviness G, Raymond E, Ganesan R, et al. Structural and functional attributes of the interleukin-36 receptor. *Journal of Biological Chemistry*. 2016;291(32): 16597–16609. Available from: doi:10.1074/jbc.M116.723064

36. Sims JE, Smith DE. The IL-1 family: regulators of immunity. *Nature Reviews Immunology*. 2010;10(2): 89–102. Available from: doi:10.1038/nri2691

37. Saha SS, Singh D, Raymond EL, Ganesan R, Caviness G, Grimaldi C, et al. Signal transduction and intracellular trafficking by the interleukin 36 receptor. *Journal of Biological Chemistry*. 2015;290(39): 23997–24006. Available from: doi:10.1074/jbc.M115.653378

38. Blumberg H, Dinh H, Trueblood ES, Pretorius J, Kugler D, Weng N, et al. Opposing activities of two novel members of the IL-1 ligand family regulate skin inflammation. *The Journal of Experimental Medicine*. 2007;204(11): 2603–2614. Available from: doi:10.1084/jem.20070157

39. *Infevers IL36RN variants*. [Online] Available from: https://fmf.igh.cnrs.fr/ISSAID/infevers/search.php?n=13

40. Sugiura K, Takemoto A, Yamaguchi M, Takahashi H, Shoda Y, Mitsuma T, et al. The majority of generalized pustular psoriasis without psoriasis vulgaris is caused by deficiency of interleukin-36 receptor antagonist. *The Journal of Investigative Dermatology*. 2013;133(11): 2514–2521. Available from: doi:10.1038/jid.2013.230

41. Mössner R, Wilsmann-Theis D, Oji V, Gkogkolou P, Löhr S, Schulz P, et al. The genetic basis for most patients with pustular skin disease remains elusive. *British Journal of*

Dermatology. 2018;178(3): 740–748. Available from: doi:10.1111/bjd.15867

42. Mössner R, Frambach Y, Wilsmann-Theis D, Löhr S, Jacobi A, Weyergraf A, et al. Palmoplantar Pustular Psoriasis Is Associated with Missense Variants in CARD14, but Not with Loss-of-Function Mutations in IL36RN in European Patients. *Journal of Investigative Dermatology*. 2015;135(10): 2538–2541. Available from: doi:10.1038/jid.2015.186

43. Takahashi T, Fujimoto N, Kabuto M, Nakanishi T, Tanaka T. Mutation analysis of IL36RN gene in Japanese patients with palmoplantar pustulosis. *The Journal of Dermatology*. 2017;44(1): 80–83. Available from: doi:10.1111/1346-8138.13551

44. Setta-Kaffetzi N, Navarini AA, Patel VM, Pullabhatla V, Pink AE, Choon S-E, et al. Rare pathogenic variants in IL36RN underlie a spectrum of psoriasis-associated pustular phenotypes. *The Journal of Investigative Dermatology*. 2013;133(5): 1366–1369. Available from: doi:10.1038/jid.2012.490

45. Robinson MS. Adaptable adaptors for coated vesicles. *Trends in Cell Biology*. 2004;14(4): 167–174. Available from: doi:10.1016/j.tcb.2004.02.002

46. Setta-Kaffetzi N, Simpson MA, Navarini AA, Patel VM, Lu HC, Allen MH, et al. AP1S3 mutations are associated with pustular psoriasis and impaired toll-like receptor 3 trafficking. *The American Journal of Human Genetics*. 2014;94(5): 790–797. Available from: doi:10.1016/j.ajhg.2014.04.005

47. Mahil SK, Twelves S, Farkas K, Setta-Kaffetzi N, Burden AD, Gach JE, et al. AP1S3 Mutations Cause Skin Autoinflammation by Disrupting Keratinocyte Autophagy and Up-Regulating IL-36 Production. *Journal of Investigative Dermatology*. 2016;136(11): 2251–2259. Available from: doi:10.1016/j.jid.2016.06.618

48. Berki DM, Liu L, Choon SE, Burden AD, Griffiths CEM, Navarini AA, et al. Activating CARD14 mutations are associated with generalized pustular psoriasis but rarely account for familial recurrence in psoriasis vulgaris. *Journal of Investigative Dermatology*. 2015;135(12): 2964–2970. Available from: doi:10.1038/jid.2015.288

49. Lenz G, Davis RE, Ngo VN, Lam L, George TC, Wright GW, et al. Oncogenic CARD11 mutations in human diffuse large B cell lymphoma. *Science*. 2008;319(5870): 1676–1679. Available from: doi:10.1126/science.1153629

50. Sugiura K, Muto M, Akiyama M. CARD14 c.526G>C (p.Asp176His) Is a Significant Risk Factor for Generalized Pustular Psoriasis with Psoriasis Vulgaris in the Japanese Cohort. *The Journal of Investigative Dermatology*. 2014;134(6): 1–11. Available from: doi:10.1038/jid.2014.46

51. Jordan CT, Cao L, Roberson EDO, Pierson KC, Yang C-F, Joyce CE, et al. PSORS2 is due to mutations in CARD14. *The American Journal of Human Genetics*. 2012;90(5): 784–795. Available from: doi:10.1016/j.ajhg.2012.03.012

52. Fuchs-Telem D, Sarig O, Van Steensel M a M, Isakov O, Israeli S, Nousbeck J, et al. Familial pityriasis rubra pilaris is caused by mutations in CARD14. *The American Journal of Human Genetics*. 2012;91(1): 163–170. Available from: doi:10.1016/j.ajhg.2012.05.010

53. Jordan CT, Cao L, Roberson EDO, Duan S, Helms CA, Nair RP, et al. Rare and common variants in CARD14, encoding an epidermal regulator of NF-kappaB, in psoriasis. *American Journal of Human Genetics*. 2012;90(5): 796–808. Available from: doi:10.1016/j.ajhg.2012.03.013

54. Zotti T, Polvere I, Voccola S, Vito P, Stilo R. CARD14/CARMA2 Signaling and its Role in Inflammatory Skin Disorders. *Frontiers in immunology*. 2018;9: 2167. Available from:

doi:10.3389/fimmu.2018.02167

55. Van Nuffel E, Schmitt A, Afonina IS, Schulze-Osthoff K, Beyaert R, Hailfinger S. CARD14-Mediated Activation of Paracaspase MALT1 in Keratinocytes: Implications for Psoriasis. *Journal of Investigative Dermatology*. 2017;137(3): 569–575. Available from: doi:10.1016/j.jid.2016.09.031

56. Johnston A, Xing X, Wolterink L, Barnes DH, Yin ZQ, Reingold L, et al. IL-1 and IL-36 are dominant cytokines in generalized pustular psoriasis. *Journal of Allergy and Clinical Immunology*. 2017;140(1): 109–120. Available from: doi:10.1016/j.jaci.2016.08.056

57. van der Fits L, Mourits S, Voerman JSA, Kant M, Boon L, Laman JD, et al. Imiquimod-Induced Psoriasis-Like Skin Inflammation in Mice Is Mediated via the IL-23/IL-17 Axis. *The Journal of Immunology*. 2009;182(9): 5836–5845. Available from: doi:10.4049/jimmunol.0802999

58. Tortola L, Rosenwald E, Abel B, Blumberg H, Schäfer M, Coyle AJ, et al. Psoriasiform dermatitis is driven by IL-36-mediated DC-keratinocyte crosstalk. *Journal of Clinical Investigation*. 2012;122(11): 3965–3976. Available from: doi:10.1172/JCI63451

59. Blumberg H, Dinh H, Dean C, Trueblood ES, Bailey K, Shows D, et al. IL-1RL2 and Its Ligands Contribute to the Cytokine Network in Psoriasis. *The Journal of Immunology*. 2010;185(7): 4354–4362. Available from: doi:10.4049/jimmunol.1000313

60. Goodwin S, McPherson JD, McCombie WR. Coming of age: Ten years of next-generation sequencing technologies. *Nature Reviews Genetics*. 2016;17(6): 333–351. Available from: doi:10.1038/nrg.2016.49

61. Heather JM, Chain B. The sequence of sequencers: The history of sequencing DNA.

Genomics. 2016; Available from: doi:10.1016/j.ygeno.2015.11.003

62. Bentley DR, Balasubramanian S, Swerdlow HP, Smith GP, Milton J, Brown CG, et al. Accurate whole human genome sequencing using reversible terminator chemistry. *Nature*. 2008; Available from: doi:10.1038/nature07517

63. Rothberg JM, Hinz W, Rearick TM, Schultz J, Mileski W, Davey M, et al. An integrated semiconductor device enabling non-optical genome sequencing. *Nature*. 2011;475(7356): 348–352. Available from: doi:10.1038/nature10242

64. Valouev A, Ichikawa J, Tonthat T, Stuart J, Ranade S, Peckham H, et al. A high-resolution, nucleosome position map of C. elegans reveals a lack of universal sequence-dictated positioning. *Genome Research*. 2008;18(7): 1051–1063. Available from: doi:10.1101/gr.076463.108

65. Drmanac R, Sparks AB, Callow MJ, Halpern AL, Burns NL, Kermani BG, et al. Human genome sequencing using unchained base reads on self-assembling DNA nanoarrays. *Science*. 2010;327(5961): 78–81. Available from: doi:10.1126/science.1181498

66. van Dijk EL, Jaszczyszyn Y, Naquin D, Thermes C. The Third Revolution in Sequencing Technology. *Trends in Genetics*. 2018;0(0). Available from: doi:10.1016/j.tig.2018.05.008

67. Eid J, Fehr A, Gray J, Luong K, Lyle J, Otto G, et al. Real-time DNA sequencing from single polymerase molecules. *Science*. 2009;323(5910): 133–138. Available from: doi:10.1126/science.1162986

68. Clarke J, Wu HC, Jayasinghe L, Patel A, Reid S, Bayley H. Continuous base identification for single-molecule nanopore DNA sequencing. *Nature Nanotechnology*. 2009;4(4): 265–270. Available from: doi:10.1038/nnano.2009.12

69. Park PJ. ChIP–seq: advantages and challenges of a maturing technology. *Nature Reviews Genetics*. 2009;10(10): 669–680. Available from: doi:10.1038/nrg2641

70. Buenrostro JD, Wu B, Chang HY, Greenleaf WJ. ATAC-seq: A method for assaying chromatin accessibility genome-wide. *Current Protocols in Molecular Biology*. 2015;2015(1): 21.29.1-21.29.9. Available from: doi:10.1002/0471142727.mb2129s109

71. Kukurba KR, Montgomery SB. RNA sequencing and analysis. *Cold Spring Harbor Protocols*. 2015;2015(11): 951–969. Available from: doi:10.1101/pdb.top084970

72. The 1000 Genomes Project Consortium. A map of human genome variation from population scale sequencing. *Nature*. 2010;476(7319): 1061–1073. Available from: doi:10.1038/nature09534.A

73. Lek M, Karczewski KJ, Minikel E V., Samocha KE, Banks E, Fennell T, et al. Analysis of protein-coding genetic variation in 60,706 humans. *Nature*. 2016;536(7616): 285–291. Available from: doi:10.1038/nature19057

74. Kiezun A, Garimella K, Do R, Stitziel NO, Neale BM, McLaren PJ, et al. Exome sequencing and the genetic basis of complex traits. *Nature Genetics*. 2012;44(6): 623–630. Available from: doi:10.1038/ng.2303

75. Bao R, Huang L, Andrade J, Tan W, Kibbe WA, Jiang H, et al. Review of Current Methods, Applications, and Data Management for the Bioinformatics Analysis of Whole Exome Sequencing. *Libertas Academica*. 2014;13(Suppl 2): 67–82. Available from: doi:10.4137/CIN.S13779.Received

76. Olson ND, Lund SP, Colman RE, Foster JT, Sahl JW, Schupp JM, et al. Best practices for evaluating single nucleotide variant calling methods for microbial genomics. *Frontiers in*

Genetics. 2015;6(JUL): 235. Available from: doi:10.3389/fgene.2015.00235

77. *Phred-scaled Quality Scores*. [Online] Available from: https://gatkforums.broadinstitute.org/gatk/discussion/4260/phred-scaled-quality-scores

78. Dashti MJS, Gamieldien J. A practical guide to filtering and prioritizing genetic variants. *BioTechniques*. 2017;62(1): 18–30. Available from: doi:10.2144/000114492

79. Huang T, Shu Y, Cai Y-D. Genetic differences among ethnic groups. *BMC Genomics*. 2015;16(1): 1093. Available from: doi:10.1186/s12864-015-2328-0

80. Maffucci P, Bigio B, Rapaport F, Cobat A, Borghesi A, Lopez M, et al. Blacklisting variants common in private cohorts but not in public databases optimizes human exome analysis. *Proceedings of the National Academy of Sciences*. 2018;116(3): 950–959. Available from: doi:10.1073/pnas.1808403116

81. Cummings BB, Marshall JL, Tukiainen T, Lek M, Donkervoort S, Foley AR, et al. Improving genetic diagnosis in Mendelian disease with transcriptome sequencing. *Science Translational Medicine*. 2017;9(386): eaal5209. Available from: doi:10.1126/scitranslmed.aal5209

82. MacArthur DG, Balasubramanian S, Frankish A, Huang N, Morris J, Walter K, et al. A systematic survey of loss-of-function variants in human protein-coding genes. *Science*. 2012;335(6070): 823–828. Available from: doi:10.1126/science.1215040

83. Matreyek KA, Starita LM, Stephany JJ, Martin B, Chiasson MA, Gray VE, et al. Multiplex assessment of protein variant abundance by massively parallel sequencing. *Nature Genetics*. 2018;50(6): 874–882. Available from: doi:10.1038/s41588-018-0122-z

84. Sundaram L, Gao H, Padigepati SR, McRae JF, Li Y, Kosmicki JA, et al. Predicting the clinical impact of human mutation with deep neural networks. *Nature Genetics*. 2018;50(8): 1161–1170. Available from: doi:10.1038/s41588-018-0167-z

85. Omoyinmi E, Standing A, Keylock A, Price-Kuehne F, Melo Gomes S, Rowczenio D, et al. Clinical impact of a targeted next-generation sequencing gene panel for autoinflammation and vasculitis. Wang J (ed.) *PLoS ONE*. 2017;12(7): e0181874. Available from: doi:10.1371/journal.pone.0181874

86. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genetics in Medicine*. 2015;17(5): 405–423. Available from: doi:10.1038/gim.2015.30

87. Li Q, Wang K, McPherson JD, Lyon GJ, Wang K, Quintáns B, et al. InterVar: Clinical Interpretation of Genetic Variants by the 2015 ACMG-AMP Guidelines. *The American Journal of Human Genetics*. 2017;100(2): 267–280. Available from: doi:10.1016/j.ajhg.2017.01.004

88. Lim KH, Fairbrother WG. Spliceman-A computational web server that predicts sequence variations in pre-mRNA splicing. *Bioinformatics*. 2012;28(7): 1031–1032. Available from: doi:10.1093/bioinformatics/bts074

89. Yeo G, Burge CB. Maximum entropy modeling of short sequence motifs with applications to RNA splicing signals. *Journal of Computational Biology*. 2004;11(2–3): 377–394. Available from: doi:10.1089/1066527041410418

90. Desmet FO, Hamroun D, Lalande M, Collod-Bëroud G, Claustres M, Béroud C. Human

Splicing Finder: An online bioinformatics tool to predict splicing signals. *Nucleic Acids Research*. 2009;37(9): e67–e67. Available from: doi:10.1093/nar/gkp215

91. Kumar P, Henikoff S, Ng PC. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nature Protocols*. 2009;4(7): 1073–1081. Available from: doi:10.1038/nprot.2009.86

92. Choi Y, Sims GE, Murphy S, Miller JR, Chan AP. Predicting the Functional Effect of Amino Acid Substitutions and Indels. *PLoS ONE*. 2012;7(10). Available from: doi:10.1371/journal.pone.0046688

93. Adzhubei I, Jordan DM, Sunyaev SR. Predicting functional effect of human missense mutations using PolyPhen-2. *Current Protocols in Human Genetics*. 2013;Chapter 7. Available from: doi:10.1002/0471142905.hg0720s76

94. Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, et al. A method and server for predicting damaging missense mutations. *Nature Methods*. 2010;7(4): 248–249. Available from: doi:10.1038/nmeth0410-248

95. Schwarz JM, Rödelsperger C, Schuelke M, Seelow D. MutationTaster evaluates disease-causing potential of sequence alterations. *Nature Methods*. 2010;7(8): 575–576. Available from: doi:10.1038/nmeth0810-575

96. Kircher M, Witten DM, Jain P, O'Roak BJ, Cooper GM, Shendure J. A general framework for estimating the relative pathogenicity of human genetic variants. *Nature Genetics*. 2014;46(3): 310–315. Available from: doi:10.1038/ng.2892

97. Grimm DG, Azencott CA, Aicheler F, Gieraths U, Macarthur DG, Samocha KE, et al. The evaluation of tools used to predict the impact of missense variants is hindered by two

types of circularity. *Human Mutation*. 2015;36(5): 513–523. Available from: doi:10.1002/humu.22768

98. Bean LJH, Hegde MR. Clinical implications and considerations for evaluation of in silico algorithms for use with ACMG/AMP clinical variant interpretation guidelines. *Genome Medicine*. 2017;9(1): 225. Available from: doi:10.1186/s13073-017-0508-z

99. Itan Y, Shang L, Boisson B, Ciancanelli MJ, Markle JG, Martinez-Barricarte R, et al. The mutation significance cutoff: gene-level thresholds for variant predictions. *Nature Methods*. 2016;13(2): 109–110. Available from: doi:10.1038/nmeth.3739

100. Guidugli L, Shimelis H, Masica DL, Pankratz VS, Lipton GB, Singh N, et al. Assessment of the Clinical Relevance of BRCA2 Missense Variants by Functional and Computational Approaches. *The American Journal of Human Genetics*. 2018;102(2): 233–248. Available from: doi:10.1016/j.ajhg.2017.12.013

101. Majithia AR, Tsuda B, Agostini M, Gnanapradeepan K, Rice R, Peloso G, et al. Prospective functional classification of all possible missense variants in PPARG. *Nature Genetics*. 2016;48(12): 1570–1575. Available from: doi:10.1038/ng.3700

102. Walter K, Min JL, Huang J, Crooks L, Memari Y, McCarthy S, et al. The UK10K project identifies rare variants in health and disease. *Nature*. 2015;526(7571): 82–90. Available from: doi:10.1038/nature14962

103. Wong L-P, Ong RT-H, Poh W-T, Liu X, Chen P, Li R, et al. Deep whole-genome sequencing of 100 southeast Asian Malays. *The American Journal of Human Genetics*. 2013;92(1): 52–66. Available from: doi:10.1016/j.ajhg.2012.12.005

104. Koressaar T, Remm M. Enhancements and modifications of primer design program

Primer3. *Bioinformatics*. 2007;23(10): 1289–1291. Available from: doi:10.1093/bioinformatics/btm091

105. Untergasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, Remm M, et al. Primer3-new capabilities and interfaces. *Nucleic Acids Research*. 2012;40(15): e115. Available from: doi:10.1093/nar/gks596

106. Zerbino DR, Achuthan P, Akanni W, Amode MR, Barrell D, Bhai J, et al. Ensembl 2018. *Nucleic Acids Research*. 2018;46(D1): D754–D761. Available from: doi:10.1093/nar/gkx1098

107. Ye J, Coulouris G, Zaretskaya I, Cutcutache I, Rozen S, Madden TL. Primer-BLAST: a tool to design target-specific primers for polymerase chain reaction. *BMC Bioinformatics*. 2012;13(1): 134. Available from: doi:10.1186/1471-2105-13-134

108. Notredame C, Higgins DG, Heringa J. T-coffee: A novel method for fast and accurate multiple sequence alignment. *Journal of Molecular Biology*. 2000;302(1): 205–217. Available from: doi:10.1006/jmbi.2000.4042

109. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods*. 2001;25(4): 402–408. Available from: doi:10.1006/meth.2001.1262

110. *BioMol.net Extinction coefficient calculator*. [Online] Available from: http://www.biomol.net/en/tools/proteinextinction.htm

111. Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, et al. Fiji: An open-source platform for biological-image analysis. *Nature Methods*. 2012;9(7): 676–682. Available from: doi:10.1038/nmeth.2019

112. Rueden CT, Schindelin J, Hiner MC, DeZonia BE, Walter AE, Arena ET, et al. ImageJ2: ImageJ for the next generation of scientific image data. *BMC Bioinformatics*. 2017;18(1): 529. Available from: doi:10.1186/s12859-017-1934-z

113. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics*. 2009;25(16): 2078–2079. Available from: doi:10.1093/bioinformatics/btp352

114. Quinlan AR, Hall IM. BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics*. 2010;26(6): 841–842. Available from: doi:10.1093/bioinformatics/btq033

115. Wang K, Li M, Hakonarson H. ANNOVAR: Functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Research*. 2010;38(16): e164–e164. Available from: doi:10.1093/nar/gkq603

116. Liu X, Jian X, Boerwinkle E. dbNSFP: A lightweight database of human nonsynonymous SNPs and their functional predictions. *Human Mutation*. 2011;32(8): 894–899. Available from: doi:10.1002/humu.21517

117. Liu X, Wu C, Li C, Boerwinkle E. dbNSFP v3.0: A One-Stop Database of Functional Predictions and Annotations for Human Non-synonymous and Splice Site SNVs. *Human Mutation*. 2015; Available from: doi:10.1002/humu.22932

118. Yang H, Wang K. Genomic variant annotation and prioritization with ANNOVAR and wANNOVAR. *Nature Protocols*. 2015;10(10): 1556–1566. Available from: doi:10.1038/nprot.2015.105

119. *NHLBI GO Exome Sequencing Project (ESP) Exome Variant Server*. [Online] Available from:

http://evs.gs.washington.edu/EVS/

120. Shyr C, Tarailo-Graovac M, Gottlieb M, Lee JJY, Van Karnebeek C, Wasserman WW. FLAGS, frequently mutated genes in public exomes. *BMC Medical Genomics*. 2014;7(1): 1–14. Available from: doi:10.1186/s12920-014-0064-y

121. Robinson JT, Thorvaldsdóttir H, Winckler W, Guttman M, Lander ES, Getz G, et al. Integrative genomics viewer. *Nature Biotechnology*. 2011;29(1): 24–26. Available from: doi:10.1038/nbt.1754

122. Borenstein, M., Hedges, L., Higgins, J., & Rothstein H. *Comprehensive Meta-Analysis Version 3*. 2013.

123. R Core Team. *R: A language and environment for statistical computing.* [Online] R Foundation for Statistical Computing, Vienna, Austria. 2017. Available from: doi:ISBN 3-900051-07-0

124. Ogle DH. *FSA: Fisheries Stock Analysis. R package version 0.8.17.* 2017.

125. Purcell S, Cherny SS, Sham PC. Genetic power calculator: Design of linkage and association genetic mapping studies of complex traits. *Bioinformatics*. 2003;19(1): 149–150. Available from: doi:10.1093/bioinformatics/19.1.149

126. O'Doherty CJ, Macintyre C. Palmoplantar pustulosis and smoking. *British Medical Journal (Clinical research ed.)*. 1985;291(6499): 861–864. Available from: doi:10.1136/bmj.291.6499.861

127. Hagforsen E, Mustafa A, Lefvert AK, Nordlind K, Michaëlsson G. Palmoplantar pustulosis: An autoimmune disease precipitated by smoking? *Acta Dermato-Venereologica*. 2002;82(5): 341–346. Available from: doi:10.1080/000155502320624069

128. Brunasso AMG, Puntoni M, Aberer W, Delfino C, Fancelli L, Massone C. Clinical and epidemiological comparison of patients affected by palmoplantar plaque psoriasis and palmoplantar pustulosis: A case series study. *British Journal of Dermatology*. 2013;168(6): 1243–1251. Available from: doi:10.1111/bjd.12223

129. Zhang MQ. Statistical features of human exons and their flanking regions. *Human Molecular Genetics*. 1998;7(5): 919–932. Available from: doi:10.1093/hmg/7.5.919

130. Li M, Han J, Lu Z, Li H, Zhu K, Cheng R, et al. Prevalent and rare mutations in IL-36RN gene in Chinese patients with generalized pustular psoriasis and psoriasis vulgaris. *The Journal of Investigative Dermatology*. 2013;133(11): 2637–2639. Available from: doi:10.1038/jid.2013.267

131. Crooks GE, Hon G, Chandonia JM, Brenner SE. WebLogo: A sequence logo generator. *Genome Research*. 2004;14(6): 1188–1190. Available from: doi:10.1101/gr.849004

132. *WebLogo 3*. [Online] Available from: http://weblogo.threeplusone.com

133. Farooq M, Nakai H, Fujimoto A, Fujikawa H, Matsuyama A, Kariya N, et al. Mutation Analysis of the IL36RN Gene in 14 Japanese Patients with Generalized Pustular Psoriasis. *Human Mutation*. 2013;34(1): 176–183. Available from: doi:10.1002/humu.22203

134. Ellingford JM, Black GCM, Clayton TH, Judge M, Griffiths CEM, Warren RB. A novel mutation in IL36RN underpins childhood pustular dermatosis. *Journal of the European Academy of Dermatology and Venereology*. 2016;30(2): 302–305. Available from: doi:10.1111/jdv.13034

135. Eriksson MO, Hagforsen E, Lundin IP, Michaëlsson G. Palmoplantar pustulosis: A clinical and immunohistological study. *British Journal of Dermatology*. 1998;138(3): 390–398.

Available from: doi:10.1046/j.1365-2133.1998.02113.x

136. Jacobson DL, Gange SJ, Rose NR, Graham NMH. Epidemiology and estimated population burden of selected autoimmune diseases in the United States. *Clinical Immunology and Immunopathology*. 1997;84(3): 223–243. Available from: doi:10.1006/clin.1997.4412

137. Klein SL, Flanagan KL. Sex differences in immune responses. *Nature Reviews Immunology*. 2016;16(10): 626–638. Available from: doi:10.1038/nri.2016.90

138. Souyris M, Cenac C, Azar P, Daviaud D, Canivet A, Grunenwald S, et al. TLR7 escapes X chromosome inactivation in immune cells. *Science Immunology*. 2018;3(19): eaap8855. Available from: doi:10.1126/sciimmunol.aap8855

139. Liang Y, Tsoi LC, Xing X, Beamer MA, Swindell WR, Sarkar MK, et al. A gene network regulated by the transcription factor VGLL3 as a promoter of sex-biased autoimmune diseases. *Nature Immunology*. 2017;18(2): 152–160. Available from: doi:10.1038/ni.3643

140. Hagforsen E, Edvinsson M, Nordlind K, Michaëlsson G. Expression of nicotinic receptors in the skin of patients with palmoplantar pustulosis. *British Journal of Dermatology*. 2002;146(3): 383–391. Available from: doi:10.1046/j.1365-2133.2002.04640.x

141. Di Meglio P, Duarte JH, Ahlfors H, Owens NDL, Li Y, Villanova F, et al. Activation of the aryl hydrocarbon receptor dampens the severity of inflammatory skin conditions. *Immunity*. 2014;40(6): 989–1001. Available from: doi:10.1016/j.immuni.2014.04.019

142. Lee KWK, Pausova Z. Cigarette smoking and DNA methylation. *Frontiers in Genetics*. 2013;4: 132. Available from: doi:10.3389/fgene.2013.00132

143. Monick MM, Beach SRH, Plume J, Sears R, Gerrard M, Brody GH, et al. Coordinated changes in AHRR methylation in lymphoblasts and pulmonary macrophages from

smokers. *American Journal of Medical Genetics, Part B: Neuropsychiatric Genetics*. 2012;159 B(2): 141–151. Available from: doi:10.1002/ajmg.b.32021

144. Liang Y, Xing X, Beamer MA, Swindell WR, Sarkar MK, Roberts LW, et al. Six-transmembrane epithelial antigens of the prostate comprise a novel inflammatory nexus in patients with pustular skin disorders. *Journal of Allergy and Clinical Immunology*. 2017;139(4): 1217–1227. Available from: doi:10.1016/j.jaci.2016.10.021

145. *Domains of CARD14*. [Online] Available from: https://www.uniprot.org/uniprot/Q9BXL6#family_and_domains

146. Lee HJ, Zheng JJ. PDZ domains and their binding partners: Structure, specificity, and modification. *Cell Communication and Signaling*. 2010;8(1): 8. Available from: doi:10.1186/1478-811X-8-8

147. Bertin J, Wang L, Guo Y, Jacobson MD, Poyet JL, Srinivasula SM, et al. CARD11 and CARD14 Are Novel Caspase Recruitment Domain (CARD)/Membrane-associated Guanylate Kinase (MAGUK) Family Members that Interact with BCL10 and Activate NF-κB. *Journal of Biological Chemistry*. 2001;276(15): 11877–11882. Available from: doi:10.1074/jbc.M010512200

148. *European Rare And Severe Psoriasis Expert Network*. [Online] Available from: http://eraspen.eu/

149. Bal E, Lim AC, Shen M, Douangpanya J, Madrange M, Gazah R, et al. Mutation in IL36RN impairs the processing and regulatory function of the interleukin-36-receptor antagonist and is associated with DITRA syndrome. *Experimental Dermatology*. 2017; Available from: doi:10.1111/exd.13387

150. Sugiura K, Takeichi T, Kono M, Ogawa Y, Shimoyama Y, Muro Y, et al. A novel IL36RN/IL1F5 homozygous nonsense mutation, p.Arg10X, in a Japanese patient with adult-onset generalized pustular psoriasis. *British Journal of Dermatology*. 2012;167(3): 699–701. Available from: doi:10.1111/j.1365-2133.2012.10953.x

151. Tauber M, Bal E, Pei X-Y, Madrange M, Khelil A, Sahel H, et al. IL36RN Mutations Affect Protein Expression and Function: A Basis for Genotype-Phenotype Correlation in Pustular Diseases. *Journal of Investigative Dermatology*. 2016;136(9): 1811–1819. Available from: doi:10.1016/j.jid.2016.04.038

152. Körber A, Mössner R, Renner R, Sticht H, Wilsmann-Theis D, Schulz P, et al. Mutations in IL36RN in patients with generalized pustular psoriasis. *Journal of Investigative Dermatology*. 2013;133(11): 2634–2637. Available from: doi:10.1038/jid.2013.214

153. Takeichi T, Togawa Y, Okuno Y, Taniguchi R, Kono M, Matsue H, et al. A newly revealed IL36RN mutation in sibling cases complements our IL36RN mutation statistics for generalized pustular psoriasis. *Journal of Dermatological Science*. 2017;85(1): 58–60. Available from: doi:10.1016/j.jdermsci.2016.10.009

154. Hayashi M, Nakayama T, Hirota T, Saeki H, Nobeyama Y, Ito T, et al. Novel IL36RN gene mutation revealed by analysis of 8 Japanese patients with generalized pustular psoriasis. *Journal of Dermatological Science*. 2014;76(3): 267–269. Available from: doi:10.1016/j.jdermsci.2014.10.008

155. Kanazawa N, Nakamura T, Mikita N, Furukawa F. Novel IL36RN mutation in a Japanese case of early onset generalized pustular psoriasis. *Journal of Dermatology*. 2013;40(9): 749–751. Available from: doi:10.1111/1346-8138.12227

156. Arostegui J, Vicente-Villa M, Chaves A, Gonzalez-Roca E, Ruiz-Ortiz E, Rius J, et al. P02-

002 - IL36RN mutations in patients with DITRA. *Pediatric Rheumatology*. 2013;11(Suppl 1): A109. Available from: doi:10.1186/1546-0096-11-S1-A109

157. Towne JE, Renshaw BR, Douangpanya J, Lipsky BP, Shen M, Gabel CA, et al. Interleukin-36 (IL-36) ligands require processing for full agonist (IL-36α, IL-36β, and IL-36γ) or antagonist (IL-36Ra) activity. *Journal of Biological Chemistry*. 2011;286(49): 42594–42602. Available from: doi:10.1074/jbc.M111.267922

158. Han X, Chen S, Flynn E, Wu S, Wintner D, Shen Y. Distinct epigenomic patterns are associated with haploinsufficiency and predict risk genes of developmental disorders. *Nature Communications*. 2018;9(1): 2138. Available from: doi:10.1038/s41467-018-04552-7

159. *Exome Aggregation Consortium (ExAC)*. [Online] Available from: http://exac.broadinstitute.org

160. Harden JL, Lewis SM, Pierson KC, Suárez-Fariñas M, Lentini T, Ortenzio FS, et al. CARD14 Expression in Dermal Endothelial Cells in Psoriasis. *PLoS ONE*. 2014;9(11): e111255. Available from: doi:10.1371/journal.pone.0111255

161. Johnston A, Xing X, Guzman AM, Riblett M, Candace M, Ward NL, et al. IL-1F5, -F6, -F8, and -F9: a novel IL-1 family signaling system that is active in psoriasis and promotes keratinocyte antimicrobial peptide expression. *Journal of Immunology*. 2012;186(4): 2613–2622. Available from: doi:10.4049/jimmunol.1003162.IL-1F5

162. Lee S, Abecasis GR, Boehnke M, Lin X. Rare-Variant Association Analysis: Study Designs and Statistical Tests. *The American Journal of Human Genetics*. 2014;95(1): 5–23. Available from: doi:10.1016/j.ajhg.2014.06.009

163. Hajioff S, McKee M. The health of the Roma people: a review of the published literature. *Journal of Epidemiology and Community Health*. 2000;54(11): 864–869. Available from: doi:10.1136/jech.54.11.864

164. Thomas JD, Doucette MM, Thomas DC, Stoeckle JD. Disease, lifestyle, and consanguinity in 58 American Gypsies. *Lancet*. 1987;2(8555): 377–379. Available from: doi:10.1016/S0140-6736(87)92392-0

165. Rehder CW, David KL, Hirsch B, Toriello H V., Wilson CM, Kearney HM. American College of medical genetics and genomics: Standards and guidelines for documenting suspected consanguinity as an incidental finding of genomic testing. *Genetics in Medicine*. 2013;15(2): 150–152. Available from: doi:10.1038/gim.2012.169

166. Wu D, Dou J, Chai X, Bellis C, Wilm A, Shih CC, et al. Large-scale whole-genome sequencing of three diverse Asian populations in Singapore. *bioRxiv*. 2018; 390070. Available from: doi:10.1101/390070

167. Purcell S. *Genetic Power Calculator*. [Online] Available from: http://zzz.bwh.harvard.edu/gpc/cc2.html

168. Ren X, Farías GG, Canagarajah BJ, Bonifacino JS, Hurley JH. Structural Basis for Recruitment and Activation of the AP-1 Clathrin Adaptor Complex by Arf1. *Cell*. 2013;152(4): 755–767. Available from: doi:10.1016/j.cell.2012.12.042

169. *Ensembl ZNF33A orthologues*. [Online] Available from: http://www.ensembl.org/Homo_sapiens/Gene/Compara_Ortholog?g=ENSG00000189180;r=10:38010650-38065088

170. Urrutia R. KRAB-containing zinc-finger repressor proteins. *Genome Biology*. 2003;4(10):

231. Available from: doi:10.1186/gb-2003-4-10-231

171. Lupo A, Cesaro E, Montano G, Zurlo D, Izzo P, Costanzo P. KRAB-Zinc Finger Proteins: A Repressor Family Displaying Multiple Biological Functions. *Current Genomics*. 2013;14(4): 268–278. Available from: doi:10.2174/1389202911314999000 2

172. Hamed SN, Sanie M, Frank WS, Michael G, Kathy NL, Ally Y, et al. C2H2 zinc finger proteins greatly expand the human regulatory lexicon. *Nat Biotechnol*. 2015;33(5): 555–562. Available from: doi:10.1038/nbt.3128

173. The UniProt Consortium. UniProt: a hub for protein information. *Nucleic Acids Research*. 2014;43(D1): D204-212. Available from: doi:10.1093/nar/gku989

174. *ZNF33A protein domains*. [Online] Available from: https://www.uniprot.org/uniprot/Q06730#family_and_domains

175. Zumbusch LR. Psoriasis und Pustuloses Exanthem. *Arch Dermatol Syphilol.* 1910;99: 335–346.

176. Ellis AG, Flohr C, Drucker AM. Network meta-analyses of systemic treatments for psoriasis: a critical appraisal. *British Journal of Dermatology*. 2018; Available from: doi:10.1111/bjd.17335

177. Capon F, Burden AD, Trembath RC, Barker JN. Psoriasis and other complex trait dermatoses: From loci to functional pathways. *Journal of Investigative Dermatology*. 2012;132(3 PART 2): 915–922. Available from: doi:10.1038/jid.2011.395

178. Landrum MJ, Lee JM, Riley GR, Jang W, Rubinstein WS, Church DM, et al. ClinVar: Public archive of relationships among sequence variation and human phenotype. *Nucleic Acids Research*. 2014;42(D1): D980-5. Available from: doi:10.1093/nar/gkt1113

179. Sarrauste de Menthière C, Terrière S, Pugnère D, Ruiz M, Demaille J, Touitou I. INFEVERS: the Registry for FMF and hereditary inflammatory disorders mutations. *Nucleic Acids Research*. 2003;31(1): 282–285. Available from: http://www.ncbi.nlm.nih.gov/pubmed/12520003

180. Thenappan T, Ormiston ML, Ryan JJ, Archer SL. Pulmonary arterial hypertension: pathogenesis and clinical management. *BMJ (Clinical research ed.)*. 2018;360: j5492. Available from: doi:10.1136/bmj.j5492

181. Chia R, Chiò A, Traynor BJ. Novel genes associated with amyotrophic lateral sclerosis: diagnostic and clinical implications. *The Lancet Neurology*. 2018;17(January): 94–102. Available from: doi:10.1016/S1474-4422(17)30401-5

182. Van Rheenen W, Shatunov A, Dekker AM, McLaughlin RL, Diekstra FP, Pulit SL, et al. Genome-wide association analyses identify new risk variants and the genetic architecture of amyotrophic lateral sclerosis. *Nature Genetics*. 2016;48(9): 1043–1048. Available from: doi:10.1038/ng.3622

183. Zhu N, Welch CL, Wang J, Allen PM, Gonzaga-Jauregui C, Ma L, et al. Rare variants in SOX17 are associated with pulmonary arterial hypertension with congenital heart disease. *Genome Medicine*. 2018;10(1): 56. Available from: doi:10.1186/s13073-018-0566-x

184. PracticeUpdate Editorial Team. *EADV 2018: The Anti–IL-36 Receptor Antibody 655130 Appears Safe and Effective in Acute Generalized Pustular Psoriasis*. [Online] PracticeUpdate. Available from: https://www.practiceupdate.com/content/eadvnbsp2018-the-antindashil-36-receptor-antibody-655130-appears-safe-and-effective-in-acute-generalized-pustular-

psoriasis/74054

185. Scharnitz T. *Focus on Psoriasis: A report from the 2018 European Academy of Dermatology and Venereology Congress in Paris, France*. [Online] 2018. Available from: https://www.psoriasiscouncil.org/ipcarticles_2018eadv.htm

Appendix I

VISIT DATE		Participant ID	

PLUM Case Report Form

REGISTRATION

Participant Name	
NHS Number	
Date of birth (dd/mm/yyyy)	/ /
Date of consent (dd/mm/yyyy)	/ /

Has consent also been given for:

	Yes	No
Recall		
Skin microbiopsies		

DEMOGRAPHICS

Sex		Male
		Female

Patient's ethnic group		White
		Asian or Asian British (Indian Subcontinent)
		Black or Black British
		Chinese, Japanese, Korean, Indochinese
		Mixed race
		Other
If "Other" please provide details		
Country of Origin		

VISIT DATE		Participant ID	

HISTORY

CLINICAL DIAGNOSIS

More than one may apply

	Yes	No	Date of onset	Date diagnosed by a dermatologist
Generalised pustular psoriasis				
Palmo-plantar pustulosis				
Acrodermatitis continua of Hallopeau				
Subacute annular pustular psoriasis				
Plaque psoriasis with pustulation within plaques				
Palmoplantar psoriasis with pustulation				
Acute generalised exanthematous pustulosis				
Chronic plaque psoriasis				
Flexural/intertriginous				
Seborrhoeic psoriasis				
Scalp psoriasis				
Psoriasis of palms/soles (non-pustular)				
Nail involvement				
Fissured tongue				
Guttate psoriasis				
Unstable psoriasis				
Psoriatic arthritis				
Reactive arthritis				
Other				
If other please provide details:				

Phenotype completed by (clinician name and signature)	

VISIT DATE		Participant ID	

FAMILY HISTORY:

Does the participant have any family history of the following?

	Yes	No	Affected relative(s)
Plaque-type psoriasis			
Pustular psoriasis			
	Yes	**No**	**Affected relative(s)**
Other psoriasis:			

Information on parental relatedness can really help genetic studies particularly in rare diseases such as pustular psoriasis but is a sensitive topic. The next question should be addressed carefully.

Phrases such as: "Are your parents related (except by marriage)?" "Are there any surnames or maiden names in common in your family?" and "Did your parents have the same surname before they were married?" may be helpful.

Are the participants' parents related to each other?		Have not asked
		Yes
		No
		Participant doesn't know
If "Yes" please describe relatedness (e.g. first cousins)		

MEDICAL HISTORY

CO-MORBIDITIES:

Does the participant have any co-morbidities? Please provide details

Co-morbidity	Year of onset

VISIT DATE		Participant ID	

TREATMENT HISTORY

Has the participant ever been treated with the following systemic therapies and what was its effect on the participants' pustular psoriasis during last use, if any?

Please provide details of the **most recent period of use including current**

	Start Date	Freq.	Dose	First Biologic?	End Date	Reason for stopping	Drug route
Etanercept							
Infliximab							
Adalimumab							
Ustekinumab							
Secukinumab							
Ixekizumab							
UVB/PUVA							
Acitretin							
Ciclosporin							
Methotrexate							

Is the participant **currently** using any topical therapies?

	Topical	Start date
1		
2		
3		
4		
5		
6		

VISIT DATE		Participant ID	

CONCOMITANT MEDICATIONS

Please list any **current** concomitant medication

Medication prescribed	Indication	Dose	Freq.	Date Started

COURSE of DISEASE

What best describes the participant's current pustular psoriasis?

	Yes	No
Persistent (ie >3 months)		
Relapsing remitting		
Single flare		

FLARES

If disease is relapsing:

Average number of flares per year	
Nature of flares (e.g generalised pustular)	
Date of last flare	

VISIT DATE		Participant ID	

Do any of the below apply during a **disease flare**? If yes, please provide the **maximum ever recorded** values if available.

	Yes	No	Not applicable	Unknown
Malaise/fatigue				
Increased CRP				
If yes, max CRP:				
Increased blood count				
If yes, max Leukocytes:			**max Neutrophils:**	
Fever				
If yes, max temperature:				

For all participants, does their pustular psoriasis get worse with any of the following:

	Yes	No	Not applicable	Unknown
Alcohol intake				
Episodes of stress				
Pregnancy				
Viral infection				
Streptococcal infection/cough/sore throat/fever				
If yes to strep infection was it confirmed by ASOT/throat swab?				
Drug (please name):				
Other (please detail):				

VISIT DATE		Participant ID	

CLINICAL ASSESSMENTS

Date assessed	/ /

Weight (kg)	
Height (cm)	

	Yes	No	Previously	Details (e.g. no. smoked per day, age started, age stopped)
Smoker?				

CURRENT DISEASE SEVERITY

Body Surface Area	
DLQI	
PGA	
PASI score	
ppPASI (if applicable)	

LOCATION OF PUSTULES (as assessed and currently visible)

	Yes	No
Head/neck		
Trunk		
Upper limbs		
Lower limbs		
Flexural areas		
Oral mucosa		

VISIT DATE		Participant ID	

	Yes	No
Palms/soles		
Nail apparatus		
Distal phalanx		
Subungual hyperkeratosis		
Permanent nail loss		
Number of digits involved		

Are any of the following available?

	Yes	No	Confirming diagnosis? Lesional? Please provide details
Clinical photos			
Histology			
Histology report			
Histology block			

RESEARCH SAMPLES

	Yes	No	Date taken
DNA (8ml) PURPLE top			

COMPLETED BY

Signed	
Print name	
Date	

Appendix II

Target	Primer ID	Sequence (5' to 3')	Anneal temp (°C)	Application
***IL36RN* Exon 2**	IL36RN_1F	GCTCCGTGGAGGCTGTTC	62	PCR and Sanger sequencing
	IL36RN_1R	CACAATTTCCCAGCTGCAAT		
***IL36RN* Exon 3**	IL36RN_2F	GGAGACAAGGCTGTGCTGTT	59	PCR and Sanger sequencing
	IL36RN_2R	GCTTAGAGCCTGGTTTGTGC		
***IL36RN* Exon 4**	IL36RN_3F	CTGCTGAGAAGCCTCCCTTC	62	PCR and Sanger sequencing
	IL36RN_3R	CAAAGCTGCCATCAACAGAA		
***IL36RN* Exon 5**	IL36RN_4F	TTCTGTTGATGGCAGCTTTG	59	PCR and Sanger sequencing
	IL36RN_4R	GGTCAGGTGCCCACTAAGTC		
***AP1S3* Exon 2**	AP1S3_2F	TTTCAGTGTCTTTGCAGAACG	59	PCR and Sanger sequencing
	AP1S3_2R	CCCCAGCCTTCAAAGATTTC		
***CARD14* Exon 3**	CARD14_3F	ACCCAGCAGAACCCAGAAA	64	PCR and Sanger sequencing
	CARD14_3R	AAGGGGGAGTAGGGCAAAT		
***CARD14* Exon 4**	CARD14_4F	TGCTCACCTGCTCACCTAC	63	PCR and Sanger sequencing
	CARD14_4R	AAGGAGTTCCAGGGAGATGG		

***ARFGAP2* Exon 1+2**	ARFGAP2 Ex1+2F	AGATCGGACTCCAATCACCC	66	PCR and Sanger sequencing
	ARFGAP2 Ex1+2R	CTTCCCACCGAATTCAGAGC		
***ARFGAP2* Exon 3**	ARFGAP2_3F	GACAGGTATCCGGGTTGC	59	PCR and Sanger sequencing
	ARFGAP2_3R	TGACCAGTTTCGAAGTTTTCAG		
***ARFGAP2* Exon 4+5**	ARFGAP2_4-5F	AACCGGTGTCAGTAGCGTGT	62	PCR and Sanger sequencing
	ARFGAP2_4-5R	CTAGGCCTACCCAGCAGGA		
***ARFGAP2* Exon 6+7**	ARFGAP2_6-7F	TGATTTCTTGTCCACAAGGTG	59	PCR and Sanger sequencing
	ARFGAP2_6-7R	CAGGCAGTAGGACCTCTGAA		
***ARFGAP2* Exon 8**	ARFGAP2_8F	CCTTGCCTGAAGCTGTTCTT	59	PCR and Sanger sequencing
	ARFGAP2_8R	GGTAAGTGGTAAGGTCAAGGAGTG		
***ARFGAP2* Exon 9+10**	ARFGAP2_9-10F	AGAAGGGCCTTTTCCTTGTG	59	PCR and Sanger sequencing
	ARFGAP2_9-10R	GCATCCCAAACTGCTAGGAA		
***ARGAP2* Exon 10**	ARFGAP2_10F_INT	ATGTAAGTGTTTGCTGCCGG	n/a	Sanger sequencing
***ARFGAP2* Exon 11**	ARFGAP2_11F	TGGGTAAGGACAGAAGGCTC	60	PCR and Sanger sequencing
	ARFGAP2_11R	TTGTCCTTGTACCTAGGGAGA		
***ARFGAP2* Exon 12**	ARFGAP2_12F	TTGGTACTTTCGCCTCTGGA	62	PCR and Sanger sequencing
	ARFGAP2_12R	GTTCTTCTAGACAGTGCTGCC		

***ARFGAP2* Exon 13**	ARFGAP2_13F	AATGCTGACGAAGCTGTGTG	62	PCR and Sanger sequencing
	ARFGAP2_13R	CTGGCCTCATACTGTGGTGA		
***ARFGAP2* Exon 14+15**	ARFGAP2_14-15F	CTGTGTGAGTCCTTGGGTCA	62	PCR
	ARFGAP2_14-15R	AAATGCTCTTAAGGCTCAGAGG		
***ARFGAP2* Exon 14**	ARFGAP2_14R_INT	CCTGTAAAACAAGAGCAGGGT	n/a	Sanger sequencing
***ARFGAP2* Exon 15**	ARFGAP2_15F_INT	ATGGATGGAGCTCACGGAG	n/a	Sanger sequencing
***ARFGAP2* Exon 16**	ARFGAP2_16F	CAGCTCTCACCGTGGACTC	62	PCR and Sanger sequencing
	ARFGAP2_163R	CAAGGGCTGGTACTGACCAT		
***ZNF33A* Exon 2**	ZNF334_Ex2F	CCATTTCTACCGCCTATTCCG	67	PCR and Sanger sequencing
	ZNF334_Ex2R	ACCGAGATGGGGACATTGTA		
***ZNF33A* Exon 3**	ZNF334_Ex3F	TCCAGCAGTGATGATAGTTCCAG	60	PCR and Sanger sequencing
	ZNF334_Ex3R	TCTTCTGGTACTCAGAGGTG		
***ZNF33A* Exon 4**	ZNF334_Ex4F	GTGTTGATTGATCACCTCTG	62	PCR and Sanger sequencing
	ZNF334_Ex4R_v2	AGGACACTGAAAGTGCTTG		
***ZNF33A* Exon 5**	ZNF334_Ex5_P1F	CATGGGGCATTTGTTTAGC	67	PCR and Sanger sequencing
	ZNF334_Ex5_P1R	CATGCTGCAAAGTCTCCTCAT		

***ZNF33A* Exon 5**	ZNF334_Ex5_P2F	GTTCTGTCAGTGTGATTCAT	56	PCR and Sanger sequencing
	ZNF334_Ex5_P2R	TTCATAGATACCCCATGAG		
***ZNF33A* Exon 5**	ZNF334_Ex5_P3F	GTTCCATCAGATATCTCCGTCAAG	62	PCR and Sanger sequencing
	ZNF334_Ex5_P3R	CTCTGGTGTACTTTAAGGTGC		
***ZNF33A* Exon 5**	ZNF334_Ex5_P4F	GAAACCCTATCAATGTAATGCG	64	PCR and Sanger sequencing
	ZNF334_Ex5_P4R	CTGTGTGTGTTCTCTGATGTAC		
***ZNF33A* Exon 5**	ZNF334_Ex5_P5F	CTCACAGTACATCAGAGAACA	64	PCR and Sanger sequencing
	ZNF334_Ex5_P5R	CTCTGATGTTGAGCAAGTTCC		
***ZNF33A* Exon 5**	ZNF334_Ex5_P6F	GAATGTGGGAAATTCTTCAGG	58	PCR and Sanger sequencing
	ZNF334_Ex5_P6R	CTTTGGAGTAACATAAGGTG		
***ZNF33A* Exon 5**	ZNF33Aint	TGTGACTTCTGGTAGAAGGC	n/a	Sanger sequencing
***IL36RN* V2F**	Mut_V2F_F	CGCGATCGCCATGTTCCTGAGTGGGGC	n/a	Plasmid mutagenesis
	Mut_V2F_R	GCCCCACTCAGGAACATGGCGATCGCG		
***IL36RN* L21P**	Mut_L21P_F	AGAAGCTGGTTATTATGCGGATAAAGCACCTTCAATGCC	n/a	Plasmid mutagenesis
	Mut_L21P_R	GGCATTGAAGGTGCTTTATCCGCATAATAACCAGCTTCT		
***IL36RN* Q25R**	Mut_Q25R_F	GCCCTCCAGCTAGAAGCCGGTTATTATGCAGATAA	n/a	Plasmid mutagenesis
	Mut_Q25R_R	TTATCTGCATAATAACCGGCTTCTAGCTGGAGGGC		

***IL36RN* E41Q**	Mut_E41Q_F	GACCACGCTGATCTGTTCACCTTTAATGACCTTCCC	n/a	Plasmid mutagenesis
	Mut_E41Q_R	GGGAAGGTCATTAAAGGTGAACAGATCAGCGTGGTC		
***IL36RN* I42N**	Mut_I42N_F	GATTGGGGACCACGCTGTTCTCTTCACCTTTAATG	n/a	Plasmid mutagenesis
	Mut_I42N_R	CATTAAAGGTGAAGAGAACAGCGTGGTCCCCAATC		
***IL36RN* V44M**	Mut_V44M_F	ACCGATTGGGGACCATGCTGATCTCTTCACC	n/a	Plasmid mutagenesis
	Mut_V44M_R	GGTGAAGAGATCAGCATGGTCCCCAATCGGT		
***IL36RN* N47S**	Mut_N47S_F	CAGCCACCGACTGGGGACCACGCTGATC	n/a	Plasmid mutagenesis
	Mut_N47S_R	GATCAGCGTGGTCCCCAGTCGGTGGCTG		
***IL36RN* A52T**	Mut_A52T_F	GGGACAGGCTGGTATCCAGCCACCGATTG	n/a	Plasmid mutagenesis
	Mut_A52T_R	CAATCGGTGGCTGGATACCAGCCTGTCCC		
***IL36RN* C70R**	Mut_C70R_F	GCCCCACCCCACGTGACAGGCACTG	n/a	Plasmid mutagenesis
	Mut_C70R_R	CAGTGCCTGTCACGTGGGGTGGGGC		
***IL36RN* T77I**	Mut_T77I_F	GGTGGGGCAGGAGCCGATTCTAACACTAGAG	n/a	Plasmid mutagenesis
	Mut_T77I_R	CTCTAGTGTTAGAATCGGCTCCTGCCCCACC		
***IL36RN* P82L**	Mut_P82L_F	GCTCCATGATGTTCACTAGCTCTAGTGTTAGAGTC	n/a	Plasmid mutagenesis
	Mut_P82L_R	GACTCTAACACTAGAGCTAGTGAACATCATGGAGC		
***IL36RN* K93E**	Mut_K93E_F	GAAGCTCTTGGATTCCTCGGCACCAAGATAGAGCT	n/a	Plasmid mutagenesis
	Mut_K93E_R	AGCTCTATCTTGGTGCCGAGGAATCCAAGAGCTTC		

***IL36RN* Y101F**	Mut_Y101_F	GTCCCGCCGGAAGAAGGTGAAGCTCTTGGA	n/a	Plasmid mutagenesis
	Mut_Y101F_R	TCCAAGAGCTTCACCTTCTTCCGGCGGGAC		
***IL36RN* R102Q**	Mut_R012Q_F	GCCCCATGTCCCGCTGGTAGAAGGTGAAG	n/a	Plasmid mutagenesis
	Mut_R102Q_R	CTTCACCTTCTACCAGCGGGACATGGGGC		
***IL36RN* R103Q**	Mut_R013Q_F	GAGCCCCATGTCCTGCCGGTAGAAGGT	n/a	Plasmid mutagenesis
	Mut_R103Q_R	ACCTTCTACCGGCAGGACATGGGGCTC		
***IL36RN* E112K**	Mut_E112K_F	GTAGGCAGCCGACTTGAAGCTGGAGGTGA	n/a	Plasmid mutagenesis
	Mut_E112K_R	TCACCTCCAGCTTCAAGTCGGCTGCCTAC		
***IL36RN* T123R**	Mut_T123R_F	GCTTCAGGCACCCTGCACAGGAACCAG	n/a	Plasmid mutagenesis
	Mut_T123R_R	CTGGTTCCTGTGCAGGGTGCCTGAAGC		
***IL36RN* E138K**	Mut_E138K_F	CCAGCCACCATTCTTGGGAAGCTGGGTGA	n/a	Plasmid mutagenesis
	Mut_E138K_R	TCACCCAGCTTCCCAAGAATGGTGGCTGG		
***IL36RN* G141D**	Mut_G141D_F	GGGGGCATTCCAGTCACCATTCTCGGG	n/a	Plasmid mutagenesis
	Mut_G141D_R	CCCGAGAATGGTGACTGGAATGCCCCC		
***IL36RN* A144T**	Mut_A144T_F	CTGTGATGGGGGTATTCCAGCCACCATTCTCG	n/a	Plasmid mutagenesis
	Mut_A144T_R	CGAGAATGGTGGCTGGAATACCCCCATCACAG		
***IL36RN* I146V**	Mut_I146V_F	GGCTGGAATGCCCCCGTCACAGACTTCTACT	n/a	Plasmid mutagenesis
	Mut_I146V_R	AGTAGAAGTCTGTGACGGGGGCATTCCAGCC		

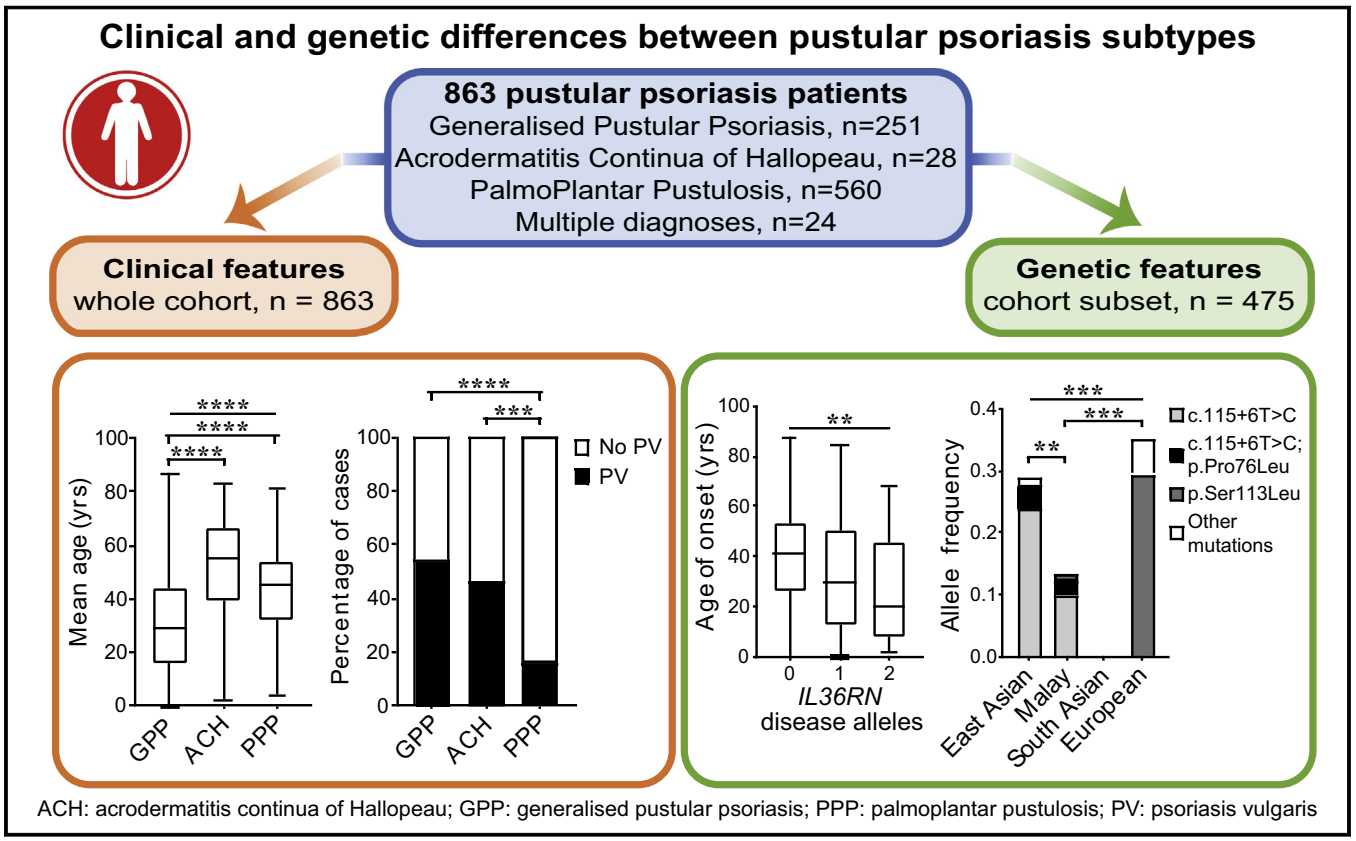
***IL36RN* Q153R**	Mut_Q153R_F	GACTTCTACTTCCAGCGGTGTGACTAGGGCAAC	n/a	Plasmid mutagenesis
	Mut_Q153R_R	GTTGCCCTAGTCACACCGCTGGAAGTAGAAGTC		
CMV	RB_CMV_F	CGGGGTCATTAGTTCATAGCC	n/a	Sanger sequencing
T7 promoter	T7F	TAATACGACTCACTATAGGG	n/a	Sanger sequencing
hGH poly(A) region	hGH rev	CAACTCAAATGTCCCACCGG	n/a	Sanger sequencing
bGH poly(A) region	BGHextra1	CTTCTAGTTGCCAGCCATC	n/a	Sanger sequencing
***ARFGAP2* full length cDNA**	ARFGAP2 Q F main	TCACTCCCCAGAGAAGAAGG	n/a	Real-time qPCR
	ARFGAP2 Q R main	CAGGCCACTGCTCTCTGTAGA		
***ARFGAP2* cDNA lacking exon 5**	ARFGAP2 Q F 2nd	ATGGCACTGATCCCCCTGC	n/a	Real-time qPCR
	ARFGAP2 Q R 2nd	TCTGTGTTGGGGCCATGCT		
***IL8* cDNA**	IL8_Fwd	TTGGCAGCCTTCCTGATTTC	n/a	Real-time qPCR
	IL8_Rev	AACTTCTCCACAACCCTCT		

Publication arising from this thesis

Clinical and genetic differences between pustular psoriasis subtypes

Sophie Twelves, MRes,[a] Alshimaa Mostafa, MD,[b,c] Nick Dand, PhD,[a] Elias Burri, MBBS,[b] Katalin Farkas, PhD,[d] Rosemary Wilson, BA,[e] Hywel L. Cooper, BM,[f] Alan D. Irvine, MD, DSc,[g] Hazel H. Oon, MD,[h] Külli Kingo, MD, PhD,[i] Sulev Köks, PhD,[j] Ulrich Mrowietz, MD,[k] Luis Puig, MD, PhD,[l] Nick Reynolds, MD,[m] Eugene Sern-Ting Tan, MBBS,[h] Adrian Tanew, MD,[n] Kaspar Torz, MD,[k] Hannes Trattner, MD,[n] Mark Valentine, MD,[o] Shyamal Wahie, MD,[p] Richard B. Warren, MB ChB, PhD,[q] Andrew Wright, MB ChB,[r] Zsuzsa Bata-Csörgő, MD,[s] Marta Szell, DSc,[t] Christopher E. M. Griffiths, MD,[q] A. David Burden, MD,[u] Siew-Eng Choon, MBBS,[v] Catherine H. Smith, MD,[e] Jonathan N. Barker, MD,[e]* Alexander A. Navarini, MD, PhD,[b]* and Francesca Capon, PhD[a]* *London, Portsmouth, Newcastle upon Tyne, Durham, Darlington, Manchester, Bradford, and Glasgow, United Kingdom; Zurich, Switzerland; Beni Suef, Egypt; Szeged, Hungary; Dublin, Ireland; Singapore; Tartu, Estonia; Kiel, Germany; Barcelona, Spain; Vienna, Austria; Seattle, Wash; and Johor Bahru, Malaysia*

GRAPHICAL ABSTRACT



ACH: acrodermatitis continua of Hallopeau; GPP: generalised pustular psoriasis; PPP: palmoplantar pustulosis; PV: psoriasis vulgaris

From [a]the Department of Medical and Molecular Genetics, School of Basic and Medical Biosciences, King's College London; [b]the Department of Dermatology, University Hospital Zurich; [c]the Department of Dermatology, Beni Suef University; [d]the Department of Medical Genetics, University of Szeged; [e]St John's Institute of Dermatology, School of Basic and Medical Biosciences, King's College London; [f]the Portsmouth Dermatology Unit, Portsmouth Hospitals Trust; [g]Paediatric Dermatology, Our Lady's Children's Hospital Crumlin, and Clinical Medicine, Trinity College Dublin; [h]the Department of Dermatology, National Skin Centre, Singapore; [i]the Department of Dermatology, University of Tartu, and the Clinic of Dermatology, Tartu University Hospital; [j]the Department of Pathophysiology, University of Tartu; [k]the Psoriasis Center at the Department of Dermatology, University Medical Center, Schleswig-Holstein, Campus Kiel; [l]the Department of Dermatology, Hospital de la Santa Creu i Sant Pau, Barcelona; [m]the Institute of Cellular Medicine, Medical School, Newcastle University and the Department of Dermatology, Royal Victoria Infirmary, Newcastle Hospitals NHS Foundation Trust, Newcastle upon Tyne; [n]the Department of Dermatology, Medical University of Vienna; [o]the Division of Dermatology, University of Washington School of Medicine, Seattle; [p]University Hospital of North Durham and Darlington Memorial Hospital; [q]the Dermatology Centre, Salford Royal Hospital, University of Manchester and the Academic Health Science Centre, Manchester; [r]St Lukes Hospital, Bradford, and the Centre for Skin Science, University of Bradford; [s]the MTA-SZTE Dermatological Research Group, Szeged, and the Department of Dermatology and Allergology, University of Szeged; [t]the MTA-SZTE Dermatological Research Group, Szeged, and the Department of Medical Genetics, University of Szeged; [u]the Institute of Infection, Inflammation and Immunity, University of Glasgow; and [v]the Department of Dermatology, Hospital Sultanah Aminah, Jeffrey Cheah School of Medicine and Health Sciences, Monash University Malaysia, Johor Bahru.

*These authors contributed equally to this work.

Background: The term pustular psoriasis indicates a group of severe skin disorders characterized by eruptions of neutrophil-filled pustules. The disease, which often manifests with concurrent psoriasis vulgaris, can have an acute systemic (generalized pustular psoriasis [GPP]) or chronic localized (palmoplantar pustulosis [PPP] and acrodermatitis continua of Hallopeau [ACH]) presentation. Although mutations have been uncovered in *IL36RN* and *AP1S3*, the rarity of the disease has hindered the study of genotype-phenotype correlations.
Objective: We sought to characterize the clinical and genetic features of pustular psoriasis through the analysis of an extended patient cohort.
Methods: We ascertained a data set of unprecedented size, including 863 unrelated patients (251 with GPP, 560 with PPP, 28 with ACH, and 24 with multiple diagnoses). We undertook mutation screening in 473 cases.
Results: Psoriasis vulgaris concurrence was lowest in PPP (15.8% vs 54.4% in GPP and 46.2% in ACH, $P < .0005$ for both), whereas the mean age of onset was earliest in GPP (31.0 vs 43.7 years in PPP and 51.8 years in ACH, $P < .0001$ for both). The percentage of female patients was greater in PPP (77.0%) than in GPP (62.5%; $P = 5.8 \times 10^{-5}$). The same applied to the prevalence of smokers (79.8% vs 28.3%, $P < 10^{-15}$). Although *AP1S3* alleles had similar frequency (0.03-0.05) across disease subtypes, *IL36RN* mutations were less common in patients with PPP (0.03) than in those with GPP (0.19) and ACH (0.16; $P = 1.9 \times 10^{-14}$ and .002, respectively). Importantly, *IL36RN* disease alleles had a dose-dependent effect on age of onset in all forms of pustular psoriasis ($P = .003$).
Conclusions: The analysis of an unparalleled resource revealed key clinical and genetic differences between patients with PPP and those with GPP. (J Allergy Clin Immunol 2018;■■■:■■■-■■■.)

Key words: *Generalized pustular psoriasis, palmoplantar pustulosis, acrodermatitis continua of Hallopeau,* IL36RN, AP1S3, *genotype-phenotype correlation*

The term pustular psoriasis refers to a group of severe inflammatory skin disorders manifesting with repeated eruptions of painful neutrophil-filled pustules. These conditions can present with acute episodes of skin pustulation and systemic upset (generalized pustular psoriasis [GPP]) or chronic pustular eruptions that affect the palms and soles (palmoplantar pustulosis [PPP]) or the tips of fingers and toes (acrodermatitis continua of Hallopeau [ACH]). Of note, all forms of the disease can be complicated by concurrent psoriasis vulgaris (PV).[1]

We and others have shown that mutations of the gene encoding the IL-36 receptor antagonist *(IL36RN)* are associated with GPP.[2,3] Although these defects are observed mostly in the homozygous or compound heterozygous state, a number of patients carrying single heterozygous changes have also been reported.[4]

Abbreviations used
ACH: Acrodermatitis continua of Hallopeau
ERASPEN: European Rare and Severe Psoriasis Expert Network
GPP: Generalized pustular psoriasis
PPP: Palmoplantar pustulosis
PV: Psoriasis vulgaris

Disease alleles associated with GPP have been identified subsequently in *AP1S3* (encoding a subunit of the adaptor protein 1 complex)[5] and *CARD14* (encoding a keratinocyte nuclear factor κB adaptor protein).[6] Of note, *IL36RN*, *CARD14*, and *AP1S3* mutations have also been described in patients with PPP and those with ACH, demonstrating a shared genetic basis for pustular forms of psoriasis.[5,7,8] Patients harboring disease alleles at 2 distinct loci (*IL36RN* and *AP1S3*; *IL36RN* and *CARD14*) have also been reported.[9,10] Thus an increasingly complex picture is emerging with evidence of substantial genetic heterogeneity, pleiotropy (the phenomenon whereby a single gene can influence more than 1 trait), and digenic inheritance.

In this context analysis of genotype-phenotype correlations would facilitate stratification of patient cohorts and streamline the genetic diagnosis of disease subtypes. However, rigorous studies have been hindered by the rarity of pustular psoriasis, which has prevented the ascertainment and standardized phenotyping of sizeable patient resources.

Here we sought to address this issue through formation of a multicenter consortium. We brought together 8 tightly phenotyped patient cohorts through a collaboration with the European Rare and Severe Psoriasis Expert Network (ERASPEN). This enabled us to ascertain a unique clinical resource, including 863 unrelated cases and exceeding by nearly 3-fold the size of any published data set. Analysis of this extended cohort revealed very significant differences in the clinical and genetic features of pustular psoriasis subtypes. Specifically, it demonstrated that PPP differs from ACH and GPP in terms of patients' demographics, disease presentation, and underlying genetic abnormalities.

METHODS

Patient ascertainment

This research was carried out in accordance with the principles of the Declaration of Helsinki and was approved by the ethics committees of

Supported by the Department of Health through the National Institute for Health Research (NIHR) BioResource Clinical Research Facility and comprehensive Biomedical Research Centre awards to Guy's and St Thomas' NHS Foundation Trust in partnership with King's College London and King's College Hospital NHS Foundation Trust (guysbrc-2012-1) and to the NIHR-Newcastle Biomedical Research Centre. This work was funded by a Medical Research Council (MRC) Stratified Medicine award (MR/L011808/1; to J.N.B., F.C., and C.H.S.) and by the Efficacy and Mechanism Evaluation (EME) Programme, an MRC and NIHR partnership (grant EME 13/50/17 to C.H.S., F.C., J.N.B., C.E.M.G., and N.R.). N.R. is also supported by the Newcastle MRC/EPSRC Molecular Pathology Node. S.T. is supported by the King's Bioscience Institute and the Guy's and St Thomas' Charity Prize PhD Programme in Biomedical and Translational Science. The European Rare and Severe Psoriasis Expert Network is funded by a PPRC grant from the European Association of Dermatology and Venereology (EADV; to A.A.N. and J.N.B.). The views expressed in this publication are those of the authors and not necessarily those of the MRC, NHS, NIHR, or Department of Health.
Received for publication February 2, 2018; revised May 14, 2018; accepted for publication June 15, 2018.
Corresponding author: Jonathan N. Barker, MD, St John's Institute of Dermatology, School of Basic and Medical Biosciences, King's College London, 9th floor Tower Wing Guy's Hospital, London SE1 9RT, United Kingdom. E-mail: jonathan.barker@kcl.ac.uk.
0091-6749

https://doi.org/10.1016/j.jaci.2018.06.038

TABLE I. Summary description of the patient cohort

	Ethnicity				Sex			Clinical diagnosis						
	European	Asian	African	Other*	Female	Male	Unknown	ACH	PPP	GPP	ACH + GPP	ACH + PPP	GPP + PPP	Total
Total	591	161	78	33	620	233	10	28	560	251	9	4	11	863

*Includes unknown ethnicity (n = 19), mixed ethnicity (n = 4), and Middle Eastern (n = 4), Finnish (n = 2), Filipino (n = 1), Hispanic (n = 1), Jamaican (n = 1), and Romani (n = 1) ethnicity.

participating institutions. Written informed consent was also obtained from all participants. The study aligned 8 patient cohorts (n = 863) recruited in the reference centers listed in Table E1 in this article's Online Repository at www.jacionline.org. The largest resource (n = 255 British and Irish cases) was provided by St John's Institute of Dermatology (London, United Kingdom) and combined a historical data set (n = 177) with patients ascertained prospectively (n = 78) through the Anakinra in Pustular Psoriasis, Response in a Controlled Trial (APRICOT) clinical trial (EudraCT no. 2015-003600-23) and its sister mechanistic study, Pustular Psoriasis, Elucidating Underlying Mechanisms (PLUM). An additional 40 affected subjects (listed as "others" in Table E1) were recruited outside the main reference centers by clinicians who sent individual samples to the ERASPEN Consortium or St John's Institute of Dermatology.

Pustular psoriasis was diagnosed by expert dermatologists based on direct clinical examination, with the ERASPEN consensus criteria[1] used in at least 506 cases. The observation of primary, sterile, macroscopically visible pustules affecting nonacral skin (GPP), palms/soles (PPP), or the nail apparatus (ACH) was the main inclusion criterion. Conversely, the occurrence of pustules restricted to the edges of psoriatic plaques represented an exclusion criterion.

Mutation screening

IL36RN, *AP1S3*, and *CARD14* mutations were screened by using Sanger sequencing in 473 patients for whom DNA was available. Primer sequences and cycling conditions have been described elsewhere.[3,5,6] Nucleotide substitutions were identified by using Sequencher 4.9 (Gene Codes, Ann Arbor, Mich). The deleterious effect of the newly identified c.115+5G>A mutation was confirmed by using Spliceman and MaxEntScan,[11,12] whereas the pathogenic potential of *CARD14* alleles was assessed with Combined Annotation Dependent Depletion (CADD).[13]

Statistics

The clinical and demographic characteristics of study participants were analyzed by using a binomial test (to establish the presence of a sex bias among patients with pustular psoriasis), the χ^2 test with the Yates correction (to analyze differences in the prevalence of PV and proportion of affected female subjects across disease types), and a Kruskal-Wallis test followed by the Dunn multiple comparison test (to analyze differences in age of onset between PPP, ACH, and GPP cases). Analysis of genetic data was based on a χ^2 test with the Yates correction (to compare the frequency of disease alleles in PPP, ACH, and GPP cases and the combined prevalence of *IL36RN* mutations across ethnic groups) and a 1-tailed Fisher exact test (for association between the *IL36RN* p.Ser113Leu allele and PPP). Genotype-phenotype correlations were investigated by implementing logistic (for PV concurrence and sex ratios) and linear (for age of onset) regression analysis with disease subtype as a covariate. All tests were implemented in R software.[14]

Patients with multiple diagnoses were excluded from all statistical analyses because they could not be assigned to a single disease group.

RESULTS

Age of onset and PV concurrence rates vary significantly among disease subtypes

As members of the ERASPEN network, we previously defined consensus criteria for the diagnosis of pustular psoriasis.[1] Here we build on this work to describe the presentation of key disease features, as observed in clinical practice. We analyzed 863 unrelated patients, the majority of whom (823/863 [95.4%]) were recruited through 6 European, 1 North African, and 1 Asian reference center (Table I and see Table E1). Of note, key patients' demographics (male/female ratios and mean age of onset for various disease types) were comparable across these cohorts (see Table E2 in this article's Online Repository at www.jacionline.org).

While patients with GPP (251/863 [29.1%]) and PPP (560/863 [64.9%]) accounted for most of the data set, the ACH sample was substantially smaller (28/863 [3.2%]), reflecting the extreme rarity of this condition. Of note, the concurrence of multiple disease forms (most notably GPP with ACH and GPP with PPP) was reported in a small percentage of affected patients (24/863 [2.8%]).

A number of comorbidities were observed, with diabetes and hypertension figuring most prominently, regardless of the patient's ethnicity (see Table E3 in this article's Online Repository at www.jacionline.org). In keeping with published associations,[15] we also found that 11 (3.9%) of 281 European patients with PPP had autoimmune thyroid disease.

Mean age of onset differed considerably across disease types and was lower in patients with GPP (31.0 ± 19.7 years) than in those with PPP (43.7 ± 14.4, $P = 9.3 \times 10^{-19}$) and those with ACH (51.8 ± 20.4, $P = 1.2 \times 10^{-7}$; Fig 1, *A*, and see Table E2). Despite these marked differences, there was substantial heterogeneity within the individual disease cohorts, with very early-onset (<10 years) and very late-onset (>70 years) cases observed in all forms of pustular psoriasis.

Although the prevalence of PV in the overall data set (29.1%) was much greater than that reported for the general population (2% to 3%), concurrence rates varied among disease variants. In particular, the frequency of PV among patients affected by PPP (15.8%) was significantly lower than that seen in the ACH (46.2%, $P = .0004$) and GPP (54.4%, $P = 2.2 \times 10^{-16}$) groups (Fig 1, *B*). Although the latter result was driven in part by a very high prevalence of PV among Malaysian patients with GPP (see Table E2), the difference remained significant ($P = .01$) when the sizeable Malaysian cohort (n = 138) was removed from the analysis. Thus our investigations have demonstrated key differences between disease subtypes, highlighting PPP as a late-onset condition with low PV concurrence.

PPP is the form of pustular psoriasis most influenced by sex and smoking status

It has been reported that female patients and smokers are at greater risk of PPP than male patients and nonsmokers.[16] Here we observed a degree of sex bias in all forms of pustular psoriasis as the female/male ratio was 1.5 in patients with ACH, 1.7 in patients with GPP, and 3.5 in patients with PPP. The distortion in sex ratios

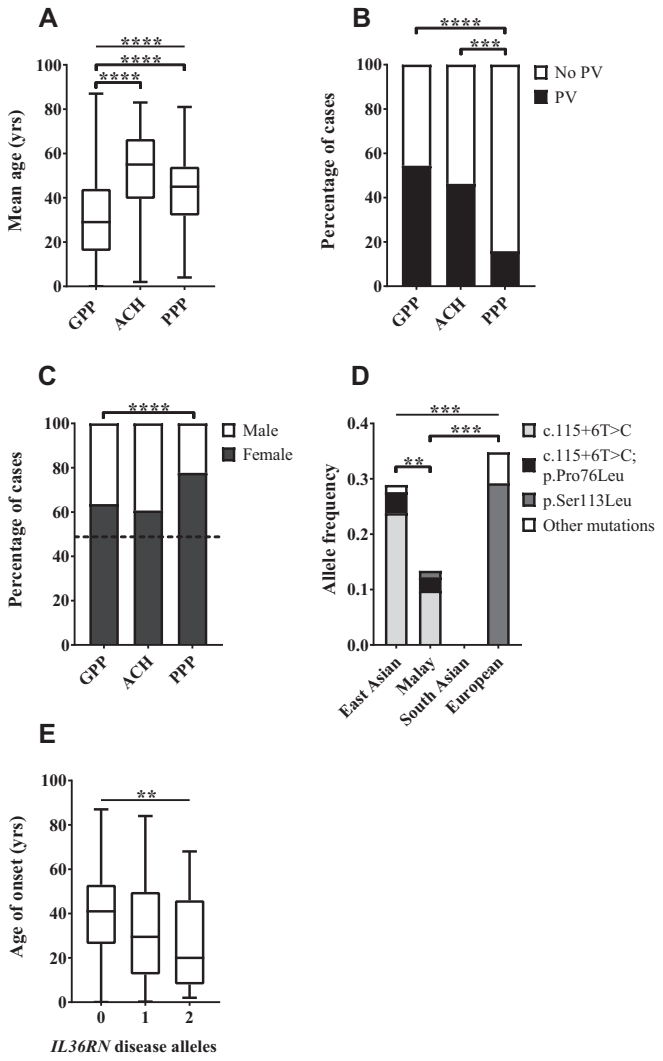


FIG 1. Features of pustular psoriasis observed in the disease cohort. **A,** Mean age of onset was compared across disease groups by using a Kruskal-Wallis test followed by the Dunn multiple comparison test. **B,** Differences in PV concurrence were analyzed with a χ^2 test. **C,** Differences in the proportion of affected female subjects were assessed by using a χ^2 test. The *dashed line* indicates the percentage of female subjects in the general population. **D,** Differences in combined frequency of *IL36RN* mutations were assessed across ethnic groups by using a χ^2 test. Pairwise comparisons were undertaken with the Fisher exact test. The analysis was restricted to patients with GPP because this is the only group for which data were available for multiple ethnicities. *Other mutations* indicates alleles seen only once in the cohort. The notation c.115+6T>C; p.Pro76Leu refers to patients carrying the 2 variants on the same haplotype. **E,** Effects of *IL36RN* mutations on age of onset were assessed by using linear regression. $^{**}P < .01$, $^{***}P < .001$, and $^{****}P < .0001$.

observed in GPP and PPP was statistically significant ($P < 10^{-5}$ and $P < 10^{-15}$, respectively) and readily recognizable in individual cohorts (see Table E2). Of note, the difference between the PPP and GPP female/male ratios was also significant ($P = 5.8 \times 10^{-5}$), highlighting PPP as the condition most influenced by sex-related factors (Fig 1, *C*).

In our data set 79.8% (249/312) of patients with PPP for whom data were available were current or past smokers. Of interest, the rate of PV concurrence was much greater in patients with PPP who smoked (or had done so in the past) compared with those who did not (12.4% vs 1.6%, $P = .009$), suggesting that cigarette smoking can modulate disease manifestations. In fact, smoking has a well-documented effect on aryl hydrocarbon receptor signaling,[17] a pathway that modulates the severity of inflammation in psoriatic skin.[18]

Although the ACH sample was too small for analysis, the percentage of smokers in the GPP data set (26/96 [28.3%]) was significantly less than that observed in patients with PPP ($P < 10^{-15}$), indicating that the adverse effect of cigarette smoking is specific to the latter condition.

Definition of a patient subset for genetic analysis

Having investigated the key clinical manifestations of pustular psoriasis, we sought to define their relationship with the patient's genotype. For this purpose, we examined the mutation status of 473 affected subjects for whom DNA was available (see Table E1). We collated genetic data previously generated by our group (n = 358)[4-6,9] while also examining 115 newly recruited cases. Importantly, Table E4 in this article's Online Repository at www.jacionline.org shows that the patient subset screened for mutations is representative of the broader data set, suggesting that the findings obtained in this sample can be generalized to the whole resource.

Frequency of *IL36RN* mutations differentiates PPP from ACH and GPP

The *IL36RN* coding sequence and exon/intron junctions were screened in the entire patient resource, uncovering 66 patients (4 with ACH, 45 with GPP, 12 with PPP, and 5 with multiple diagnoses) with disease alleles (Table II and see Table E5 in this article's Online Repository at www.jacionline.org). Thirty-six of these subjects harbored biallelic (homozygous/compound heterozygous) changes, with the remaining 30 carrying monoallelic (single heterozygous) variants. All the observed mutations had been described previously, except for a c.115+5G>A splicing variant uncovered in a North American patient with GPP (see Table E5).

IL36RN disease alleles were present in a variety of ethnic groups, with the greatest prevalence observed among patients of European (34.7%) and East Asian (28.8%) descent (Fig 1, *D*). Although we did not detect any rare changes in the 21 South Asian cases we examined, a homozygous p.Leu21Pro mutation has been described in a Pakistani GPP pedigree,[19] suggesting that deleterious *IL36RN* alleles can also be found within the Indian subcontinent.

The proportion of subjects harboring *IL36RN* disease alleles was greater in GPP and ACH (23.7% and 18.2%, respectively) compared to PPP (5.2%). Patients with GPP and those with ACH were also more likely to carry biallelic mutations compared to individuals affected by PPP (see Table E5). As a result, the prevalence of *IL36RN* mutations was significantly increased in patients with GPP (0.19) and ACH (0.16) compared with that in patients with PPP (0.03; $P = 1.9 \times 10^{-14}$ and .0018, respectively; Table II). Nonetheless, the association between *IL36RN* mutations and PPP, which has been recently questioned,[10] was statistically significant. In fact, an analysis of the recurrent p.Ser113Leu variant showed that its frequency in British patients was almost 10 times greater than that observed in population-matched control subjects ($P = 9.3 \times 10^{-8}$; odds ratio, 10.8; 95% CI, 5.3-22.0; Table III).

TABLE II. *IL36RN* and *AP1S3* mutation frequencies across disease types

	ACH	GPP	PPP	Multiple diagnoses
No. of *IL36RN*-positive patients*	4/23 (17.4%)	45/190 (23.7%)	12/234 (5.1%)	5/18 (27.8%)
IL36RN mutation count (frequency)	7/46 (0.15)	72/380 (0.19)	15/468 (0.03)	8/36 (0.22)
No. of *AP1S3*-positive patients*†	2/19 (10.5%)	4/37 (10.8%)	14/212 (6.6%)	4/11 (36.4%)
AP1S3 mutation count (frequency)	2/38 (0.05)	4/74 (0.05)	14/424 (0.03)	4/22 (0.18)

*Patients were classified as "positive" if they were carrying at least 1 mutation at the examined locus.
†p.Phe4Cys and p.Arg33Trp mutations have no frequency in East Asian populations and therefore were not screened in patients from this ethnic group.

TABLE III. Association between *IL36RN* p.Ser113Leu and PPP

	p.Ser113Leu	WT
Cases*	11 (3.6%)	291 (96.4%)
Control subjects†	26 (0.4%)	7402 (99.6%)

WT, Wild-type.
*British patients only.
†Control subjects from publicly accessible cohorts (TWINSUK and ALSPAC).

We next sought to determine whether *IL36RN* alleles were associated with key features of pustular psoriasis across disease subtypes. Therefore we implemented a regression analysis using clinical diagnosis as a covariate. Although we did not observe a consistent effect of *IL36RN* mutations on PV concurrence (see Table E6 in this article's Online Repository at www.jacionline.org), we found a significant association with early age of onset ($P = .003$; Fig 1, *E*), which was observed in all 3 forms of the disease (see Table E6). Thus *IL36RN* alleles have shared genetic effects across pustular psoriasis subtypes but occur at a very low frequency among patients with PPP.

CARD14 mutations are observed in only a small minority of cases

Although a sizeable patient subset (n = 106/473) was sequenced for the entire *CARD14* coding region, a targeted screening of exons 3 and 4 was undertaken in the rest of the sample, given that the only disease alleles associated with pustular (p.Asp176His) or plaque (p.Gly117Ser) psoriasis map to this mutation hotspot.[6,20,21]

We found 3 previously described[6] GPP patients of Chinese descent bearing the p.Asp176His variant. We did not detect any *CARD14* substitutions among European patients with GPP but observed 5 British patients with PPP harboring rare nonsynonymous changes with deleterious potential (see Table E7 in this article's Online Repository at www.jacionline.org). Although most of the above subjects (6/8 [75%]) had concurrent PV, the small size of the data set prevented us from establishing genotype-phenotype correlations.

AP1S3 mutations occur with comparable frequency across disease types

Although a substantial patient subset (n = 249) was screened for the entire coding region, the rest were sequenced only for exon 2, given that the only known *AP1S3* mutations (p.Phe4Cys, p.Arg33Trp) map to this genomic segment 2.[5,9] This revealed 24 European cases (2 patients with ACH, 4 with GPP, 14 with PPP, and 4 with multiple diagnoses) bearing the p.Phe4Cys or p.Arg33Trp changes (see Table E8 in this article's Online Repository at www.jacionline.org). No additional mutations were observed in the subjects who were screened for the entire coding region. Of note, 3 patients (2 with GPP and 1 with PPP) carried both *AP1S3* and *IL36RN* disease alleles (see Table E9 in this article's Online Repository at www.jacionline.org).

The prevalence of *AP1S3* mutations was not significantly different across disease types (Table II) and did not seem to influence PV concurrence or age of onset (see Table E10 in this article's Online Repository at www.jacionline.org). However, it was noteworthy that almost all patients with *AP1S3* disease alleles (23/24 [95.8%]) were female. Although this observation was not statistically significant ($P = .06$), a trend toward female overrepresentation was apparent in all clinical variants (see Table E10), suggesting that the penetrance of *AP1S3* mutations might be modified by sex-specific factors, such as hormone levels or X-linked modifiers.

DISCUSSION

The purpose of our study was to robustly define clinical and genetic features of pustular psoriasis by investigating a patient cohort of unprecedented size.

Initially, we sought to define the presentation of the various clinical variants through a rigorous statistical analysis of key phenotypic features. This work, which builds on the definition of consensus diagnostic criteria by ERASPEN,[1] underscores the importance of collaborative efforts in the analysis of rare diseases. Here a common case report form was used in all prospectively recruited cases, enabling standardized patient phenotyping and robust data collection. The participation of multiple centers also allowed us to monitor the effects of ascertainment bias and show that key patients' demographics were comparable across the various data sets.

Our analysis demonstrated novel and significant differences between disease subtypes. Specifically, it showed that PPP is associated with patients' demographics (very high prevalence of female subjects and smokers), clinical (low rates of PV) and genetic features (low prevalence of *IL36RN* mutations) that are clearly distinct from those observed in ACH and GPP. Given that abnormal IL-36 signaling has now been implicated in the pathogenesis of plaque psoriasis,[22] it is tempting to speculate that these observations might be correlated with each other and that the decreased prevalence of PV in PPP might be linked to the low frequency of deleterious *IL36RN* alleles in this patient group.

We also found that *IL36RN* mutations are associated with an earlier age of onset across all variants of pustular psoriasis. This validates the results we obtained originally in patients with GPP[4] and indicates that *IL36RN* should be prioritized for mutation screening when patients have disease symptoms before

the age of 30 years (40 years in the case of ACH/PPP). Given that biologics that counter the effect of *IL36RN* mutations by blocking IL-36 signaling are now under development,[23] such targeted screening could have important implications for patient management.

Our study showed that *IL36RN* mutations are the most frequent genetic abnormality observed in pustular psoriasis. In fact, deleterious *AP1S3* alleles were found in only 7% to 10% of patients, and *CARD14* variants were observed in a very small number of affected subjects. Importantly, our analysis demonstrated that known genes account only for a minority of disease cases. This is especially the case in patients with PPP, in whom the combined frequency of *AP1S3* and *IL36RN* mutations is less than 10%. Therefore additional studies will be needed to illuminate the genetic landscape of this condition, facilitate its diagnosis, and better understand the correlation between genotype and clinical phenotype. Although the discovery of novel genetic determinants has thus far been hindered by the rarity and heterogeneous nature of the disease, the ascertainment and rigorous phenotyping of our clinical resource lays a robust foundation for future gene identification studies.

We thank the Psoriasis Association for their continued support with patient recruitment. We also thank the following PLUM and APRICOT investigators for their contribution to patient ascertainment: Mahmud Ali (Worthing Hospital, Worthing, United Kingdom), Suzannah August (Poole Hospital, Poole, United Kingdom), Herve Bachelez (AP-HP Saint-Louis Hospital, Paris, France), Anthony Bewley (Whipps Cross University Hospital, London, United Kingdom), John Ingram (Cardiff University, Cardiff, United Kingdom), Susan Kelly (The Royal Shrewsbury Hospital, Shrewsbury, United Kingdom), Mohsen Korshid (Basildon Hospital, Basildon, United Kingdom), Effie Ladoyanni (Russell's Hall Hospital, Dudley, United Kingdom), and John McKenna (Leicester Royal Infirmary, Leicester, United Kingdom). We also thank the following external collaborators, who contributed to patient recruitment outside of the main reference centers: Ivona Aksentijevich (National Institutes of Health, Bethesda, Md), Sibel Dogan (Hacettepe University, Ankara, Turkey), Carlos Ferrándiz (Germans Trias i Pujol Hospital, Barcelona, Spain), Eduardo Fonseca (Complejo Hospitalario Juan Canalejo, Badalona, Spain), Joanna E. Gach (University Hospitals Coventry and Warwickshire, United Kingdom), Maja Mockenhaupt (University of Freiburg, Freiburg, Germany), Jason Pinner (Royal Prince Alfred Hospital, Sidney, Australia), Christa Prins (Geneva University Hospital, Switzerland), Annamarie Ranki (Helsinki University Central Hospital, Helsinki, Finland), Raquel Rivera (Hospital Universitario 12 de Octubre, Madrid, Spain), Marieke M Seyger (Radboud University Medical Center, Nijmegen, The Netherlands), Pere Soler-Palacin (Vall d'Hebron Research Institute, Barcelona, Spain), Eoin R. Storan (Paediatric Dermatology, Our Lady's Children's Hospital Crumlin, Dublin, Ireland), Virginia Sybert (University of Washington, Seattle, Wash), Raül Tortosa (Vall d'Hebron Hospital Research Institute, Barcelona, Spain), and Helen S. Young (the University of Manchester, Manchester, United Kingdom).

Clinical implications: The association between *IL36RN* mutations and early-onset pustular psoriasis defines a patient group that should be prioritized for *IL36RN* screening and might benefit from the development of IL-36 inhibitors.

REFERENCES

1. Navarini AA, Burden AD, Capon F, Mrowietz U, Puig L, Koks S, et al. European consensus statement on phenotypes of pustular psoriasis. J Eur Acad Dermatol Venereol 2017;31:1792-9.
2. Marrakchi S, Guigue P, Renshaw BR, Puel A, Pei XY, Fraitag S, et al. Interleukin-36-receptor antagonist deficiency and generalized pustular psoriasis. N Engl J Med 2011;365:620-8.
3. Onoufriadis A, Simpson MA, Pink AE, Di Meglio P, Smith CH, Pullabhatla V, et al. Mutations in IL36RN/IL1F5 are associated with the severe episodic inflammatory skin disease known as generalized pustular psoriasis. Am J Hum Genet 2011;89:432-7.
4. Hussain S, Berki DM, Choon SE, Burden AD, Allen MH, Arostegui JI, et al. IL36RN mutations define a severe auto-inflammatory phenotype of generalized pustular psoriasis. J Allergy Clin Immunol 2015;135:1067-70.
5. Setta-Kaffetzi N, Simpson MA, Navarini AA, Patel VM, Lu HC, Allen MH, et al. AP1S3 mutations are associated with pustular psoriasis and impaired Toll-like receptor 3 trafficking. Am J Hum Genet 2014;94:790-7.
6. Berki DM, Liu L, Choon SE, David Burden A, Griffiths CE, Navarini AA, et al. Activating CARD14 mutations are associated with generalized pustular psoriasis but rarely account for familial recurrence in psoriasis vulgaris. J Invest Dermatol 2015;135:2964-70.
7. Mossner R, Frambach Y, Wilsmann-Theis D, Lohr S, Jacobi A, Weyergraf A, et al. Palmoplantar pustular psoriasis is associated with missense variants in CARD14, but not with loss-of-function mutations in IL36RN in European patients. J Invest Dermatol 2015;135:2538-41.
8. Setta-Kaffetzi N, Navarini AA, Patel VM, Pullabhatla V, Pink AE, Choon SE, et al. Rare pathogenic variants in IL36RN underlie a spectrum of psoriasis-associated pustular phenotypes. J Invest Dermatol 2013;133:1366-9.
9. Mahil SK, Twelves S, Farkas K, Setta-Kaffetzi N, Burden AD, Gach JE, et al. AP1S3 mutations cause skin autoinflammation by disrupting keratinocyte autophagy and up-regulating IL-36 production. J Invest Dermatol 2016;136:2251-9.
10. Mössner R, Wilsmann-Theis D, Oji V, Gkogkolou P, Löhr S, Schulz P, et al. The genetic basis for most patients with pustular skin disease remains elusive. Br J Dermatol 2018;178:740-8.
11. Lim KH, Fairbrother WG. Spliceman—a computational web server that predicts sequence variations in pre-mRNA splicing. Bioinformatics 2012;28:1031-2.
12. Yeo G, Burge CB. Maximum entropy modeling of short sequence motifs with applications to RNA splicing signals. J Comput Biol 2004;11:377-94.
13. Kircher M, Witten DM, Jain P, O'Roak BJ, Cooper GM, Shendure J. A general framework for estimating the relative pathogenicity of human genetic variants. Nat Genet 2014;46:310-5.
14. R Core Team. R: a language and environment for statistical computing. Vienna: R Foundation for Statistical Computing; 2017.
15. Mrowietz U, van de Kerkhof PC. Management of palmoplantar pustulosis: do we need to change? Br J Dermatol 2011;164:942-6.
16. Burden AD, Kirby B. Psoriasis and related disorders. In: Griffiths CEM, Barker JN, Bleiker T, Chalmers RJ, Creamer D, editors. Rook's textbook of dermatology. Chichester: Wiley-Blackwell; 2016.
17. Chen Y, Widschwendter M, Teschendorff AE. Systems-epigenomics inference of transcription factor activity implicates aryl-hydrocarbon-receptor inactivation as a key event in lung cancer development. Genome Biol 2017;18:236.
18. Di Meglio P, Duarte JH, Ahlfors H, Owens ND, Li Y, Villanova F, et al. Activation of the aryl hydrocarbon receptor dampens the severity of inflammatory skin conditions. Immunity 2014;40:989-1001.
19. Ellingford JM, Black GC, Clayton TH, Judge M, Griffiths CE, Warren RB. A novel mutation in IL36RN underpins childhood pustular dermatosis. J Eur Acad Dermatol Venereol 2016;30:302-5.
20. Jordan CT, Cao L, Roberson ED, Pierson KC, Yang CF, Joyce CE, et al. PSORS2 is due to mutations in CARD14. Am J Hum Genet 2012;90:784-95.
21. Sugiura K, Muto M, Akiyama M. CARD14 c.526G>C (p.Asp176His) is a significant risk factor for generalized pustular psoriasis with psoriasis vulgaris in the Japanese cohort. J Invest Dermatol 2014;134:1755-7.
22. Mahil SK, Catapano M, Di Meglio P, Dand N, Ahlfors H, Carr IM, et al. An analysis of IL-36 signature genes and individuals with IL1RL2 knockout mutations validates IL-36 as a psoriasis therapeutic target. Sci Transl Med 2017;9:eaan2514.
23. Ganesan R, Raymond EL, Mennerich D, Woska JR, Caviness G, Grimaldi C, et al. Generation and functional characterization of anti-human and anti-mouse IL-36R antagonist monoclonal antibodies. MAbs 2017;9:1143-54.